

**CURSO DE ESPECIALIZAÇÃO EM GEOPROCESSAMENTO,  
LEVANTAMENTO E INTERPRETAÇÃO DE SOLOS  
UFRRJ - UFMG - UFV – UFRA**

**RELATÓRIO DE ANÁLISE GEOESPACIAL DE SOLOS USANDO R**

**Samuel da Costa dos Santos**

**Bragança, dezembro de 2024**

## Sumário

<b>1.</b>	<b>Resumo .....</b>	<b>3</b>
<b>2.</b>	<b>Mapa de Localização e Pontos de Amostragem.....</b>	<b>3</b>
<b>3.</b>	<b>Descrição da Base de Dados.....</b>	<b>4</b>
<b>4.</b>	<b>Estatísticas Descritivas da Variável Argila.....</b>	<b>4</b>
4.1.	Teste de Normalidade e Transformação Box-Cox .....	5
4.2.	Gráficos de Análise .....	5
4.3.	Estatísticas Descritivas Atualizadas .....	6
4.4.	Outliers na Variável Transformada.....	6
4.4.1.	Preservação da Amostragem Representativa .....	7
4.4.2.	Contexto Geoespacial e Dependência Espacial.....	7
4.4.3.	Impacto Controlado no Modelo .....	7
4.4.4.	Coerência com o Objetivo da Análise .....	7
<b>5.</b>	<b>Análise Variográfica.....</b>	<b>8</b>
5.1.	Semivariograma Experimental .....	8
5.2.	Ajuste do Modelo Teórico .....	9
5.3.	Resultados .....	10
<b>6.</b>	<b>Interpolação por Krigagem .....</b>	<b>11</b>
6.1.	Metodologia.....	11
6.1.1.	Configuração da Grade.....	11
6.1.2.	Modelo Utilizado.....	11
6.1.4.	Corte pelo Polígono Limite.....	11
6.2.	Resultados .....	11
6.2.1.	Produtos Gerados .....	12
6.3.	Considerações.....	13
<b>7.</b>	<b>Mapa de Variabilidade Espacial.....</b>	<b>14</b>
7.1.	Metodologia.....	14
7.1.1.	Cálculo do Coeficiente de Variação .....	14
7.1.2.	Representação Espacial.....	14
7.1.3.	Visualização e Escala de Cores .....	14
7.1.4.	Ajustes de Borda.....	14
7.2.	Resultados .....	15
7.3.	Considerações.....	15
<b>8.</b>	<b>Discussões.....</b>	<b>15</b>
<b>9.</b>	<b>Conclusão .....</b>	<b>16</b>
<b>10.</b>	<b>Disponibilidade de Dados e Código.....</b>	<b>17</b>
<b>11.</b>	<b>Referências .....</b>	<b>18</b>

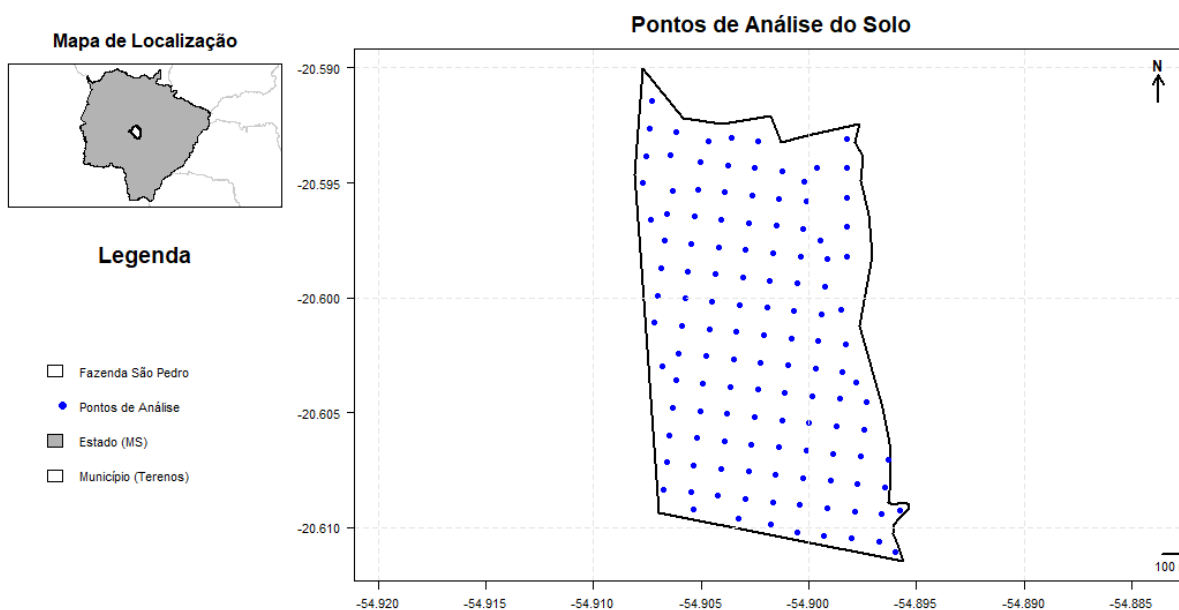
## 1. Resumo

Neste trabalho, realizamos uma análise detalhada de propriedades físico-químicas do solo em uma área específica, utilizando dados geoespaciais e ferramentas estatísticas e espaciais. O objetivo foi identificar padrões espaciais, variabilidade dos atributos analisados e apresentar informações que podem embasar práticas de manejo mais eficazes. Este relatório documenta o processo de análise, desde a preparação dos dados até a geração de mapas temáticos.

## 2. Mapa de Localização e Pontos de Amostragem

O primeiro passo foi a criação de um Mapa de Localização (Figura1), elaborado com o software R, que apresenta a delimitação do imóvel e os pontos de amostragem. O mapa foi produzido utilizando shapefiles da área de estudo (polígono do limite do imóvel) e dos pontos de análise do solo.

A área de estudo está localizada em um sistema de referência espacial uniforme (SIRGAS 2000 / UTM zona 21S), e os pontos foram distribuídos estrategicamente para cobrir toda a extensão do imóvel de maneira representativa. Esses pontos representam amostras do solo que foram analisadas para determinar propriedades específicas, como teor de argila, pH (em água e solução de  $\text{CaCl}_2$ ), entre outras.



**Figura 1: Mapa de Localização e Pontos de Amostragem**

O mapa inclui elementos de navegação, como escala gráfica, seta de norte geográfico e legenda. Esses componentes foram adicionados para facilitar a interpretação e assegurar uma visão clara da área analisada e das localizações de coleta.

### 3. Descrição da Base de Dados

A análise foi conduzida utilizando uma base de dados detalhada com informações físico-químicas do solo, composta por 31 colunas, incluindo parâmetros como pH, teores de argila, silte e areia, além de índices de saturação por bases e características químicas do solo. Cada linha representa uma amostra do terreno, totalizando 125 pontos de análise, distribuídos uniformemente em uma grade de 133 metros de distância entre os pontos nas direções leste-oeste e norte-sul. As colunas da base incluem:

<i>Amostra</i>	<i>Ca+Mgcmolc</i>	<i>Ca%</i>
<i>ID</i>	<i>Alcmolc/dm</i>	<i>Mg%</i>
<i>X</i>	<i>Al+Hcmolc/</i>	<i>K%</i>
<i>Y</i>	<i>Hcmolc/dm?</i>	<i>m%</i>
<i>pH<sub>CaCl2</sub></i>	<i>Scmolc/dm?</i>	<i>H%</i>
<i>pH<sub>Água</sub></i>	<i>Tcmolc/dm?</i>	
<i>Pmg/dm<sup>3</sup></i>	<i>V%</i>	<i>Areia%</i>
<i>MOg/dm<sup>3</sup></i>	<i>Ca/Mg</i>	<i>Silte%</i>
<i>Kcmolc/dm?</i>	<i>Ca+Mg/K</i>	<i>Argila%</i>
<i>Cacmolc/dm</i>	<i>Ca/K</i>	
<i>Mgcmolc/dm</i>	<i>Mg/K</i>	<i>Classificação</i>

Para integrar a base ao software de geoprocessamento, as colunas X e Y foram utilizadas como coordenadas em um arquivo de texto delimitado, que foi importado para o QGIS e convertido em shapefile. Este shapefile contém 125 pontos distribuídos uniformemente, representando a grade de amostragem no terreno.

### 4. Estatísticas Descritivas da Variável Argila

O parâmetro escolhido para esta análise foi o teor de argila, devido à sua relevância para a caracterização edáfica e para o manejo agrícola. Inicialmente, realizamos uma análise estatística descritiva da variável, cujos resultados foram:

- **Média:** 49,14%
- **Mediana:** 47%

- **Desvio padrão:** 4,01%
- **Mínimo:** 41%
- **Máximo:** 57%

Foi observado que a variável apresenta uma distribuição assimétrica, não atendendo aos pressupostos de normalidade. Para mitigar essa limitação, aplicamos a transformação Box-Cox, que busca estabilizar a variância e aproximar os dados de uma distribuição normal.

#### 4.1. Teste de Normalidade e Transformação Box-Cox

O teste de Shapiro-Wilk foi realizado antes e após a transformação Box-Cox. Os resultados indicaram que, apesar da transformação, os dados ainda não atingiram a normalidade, mas houve uma redução significativa da assimetria:

- **Antes da Transformação:**

##### Resultados do Teste de Normalidade

Shapiro-Wilk

Variável	Estatística W	Valor-p
Argila (Transformada)	0,92479	3,068e-06

- **Depois da Transformação:**

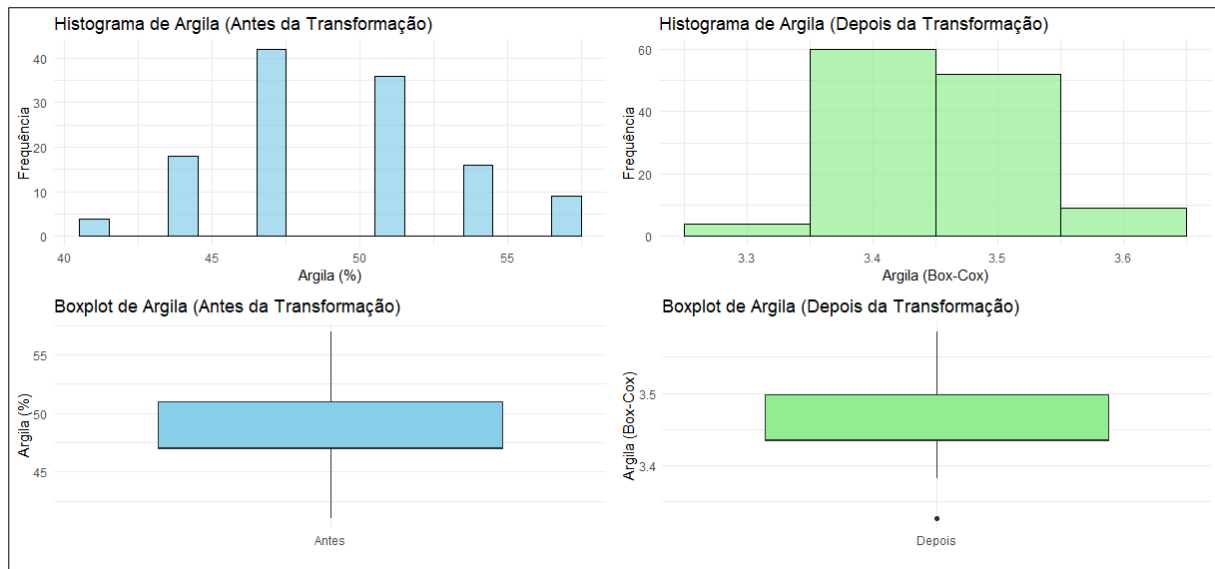
##### Resultados do Teste de Normalidade

Shapiro-Wilk

Variável	Estatística W	Valor-p
Argila (Transformada)	0,92719	4,317e-06

#### 4.2. Gráficos de Análise

Para visualizar os efeitos da transformação, foram gerados histogramas e boxplots antes e depois da aplicação do Box-Cox. Esses gráficos mostram uma redução na assimetria e na dispersão dos valores (Figura 2), mas ainda apontam para a presença de desajustes em relação à normalidade.



*Figura 2: Histogramas e Boxplots antes e depois da Transformação Box-Cox*

### 4.3. Estatísticas Descritivas Atualizadas

Após aplicar a transformação Box-Cox, as estatísticas descritivas da variável argila foram recalculadas:

- **Média:** 3,47
- **Desvio Padrão:** 0,06
- **Mediana:** 3,43
- **Mínimo:** 3,33
- **Máximo:** 3,59
- **Assimetria:** 0,0027
- **Curtose:** -0,67

A variável Argila foi normalizada parcialmente, mas não atingiu completamente a distribuição normal. Observa-se que a discrepância entre estas medidas e as originais ocorre porque a escala da variável foi modificada pela transformação Box-Cox. Assim, os valores originais, que estavam em porcentagem (e.g., 49,14%), foram transformados para uma escala baseada no parâmetro lambda da Box-Cox.

### 4.4. Outliers na Variável Transformada

Após a transformação Box-Cox, foram identificados valores que ainda se destacavam da distribuição principal, especificamente 3.325338. Esses valores foram classificados como

outliers estatísticos com base em técnicas padrão (e.g., análise de boxplot). Entretanto, optou-se por não removê-los da análise pelas seguintes razões:

#### **4.4.1. Preservação da Amostragem Representativa**

Os valores identificados como outliers remanescentes representam condições edáficas legítimas da área estudada, sendo parte integrante do processo natural de variação do solo. Segundo McBratney e Webster (1986) a remoção desses pontos poderia comprometer a representatividade da análise, especialmente em áreas onde a variabilidade é elevada e pode refletir mudanças reais nas propriedades do solo.

#### **4.4.2. Contexto Geoespacial e Dependência Espacial**

A análise geoespacial, como krigagem e variografia, baseia-se na dependência espacial dos dados. A exclusão de pontos isolados pode distorcer a continuidade espacial ou criar lacunas artificiais no modelo, prejudicando a qualidade das estimativas em locais não amostrados (ISAACS; SRIVASTAVA, 1989, p. 196). Assim, manter esses outliers ajuda a capturar variações locais relevantes.

#### **4.4.3. Impacto Controlado no Modelo**

Embora esses valores sejam classificados como outliers estatísticos, sua influência no modelo final é limitada. Osborne e Overbay (2004) explicitamente afirmam que transformações são uma forma de acomodar outliers, em vez de simplesmente removê-los. Isso é crucial para manter dados legítimos, ao mesmo tempo em que se busca atender aos pressupostos estatísticos. A transformação Box-Cox já reduziu sua magnitude e estabilizou a variância dos dados. Dessa forma, mesmo com a presença desses pontos, a distribuição geral foi significativamente ajustada para atender aos pressupostos analíticos.

#### **4.4.4. Coerência com o Objetivo da Análise**

A decisão de manter outliers em análises de solo, especialmente após transformações como a Box-Cox, está alinhada com o objetivo central de entender a variabilidade espacial e edáfica do terreno. Conforme Webster e Oliver (2007), a inclusão de todos os pontos, mesmo aqueles considerados extremos, pode ser essencial em análises geostatísticas, pois reflete a heterogeneidade intrínseca de um ambiente natural, permitindo interpretações mais próximas da realidade do campo. Nesse contexto, remover outliers pode distorcer a análise ao suprimir

variações legítimas que podem ser importantes para decisões práticas, como estratégias de manejo ou planejamento agrícola.

## 5. Análise Variográfica

Nesta etapa, foi realizada a análise variográfica para compreender a estrutura de dependência espacial dos dados da variável transformada Argila\_boxcox. O objetivo principal foi identificar o comportamento da semivariância em função da distância, permitindo o ajuste de um modelo teórico que descreva adequadamente essa relação.

### 5.1. Semivariograma Experimental

Inicialmente, o semivariograma experimental (Figura 3) foi calculado utilizando os valores transformados. Essa análise foi conduzida com o auxílio do pacote gstat, gerando pares de dados para diferentes classes de distância. O semivariograma experimental representa graficamente como a similaridade entre os pontos diminui à medida que a distância entre eles aumenta.

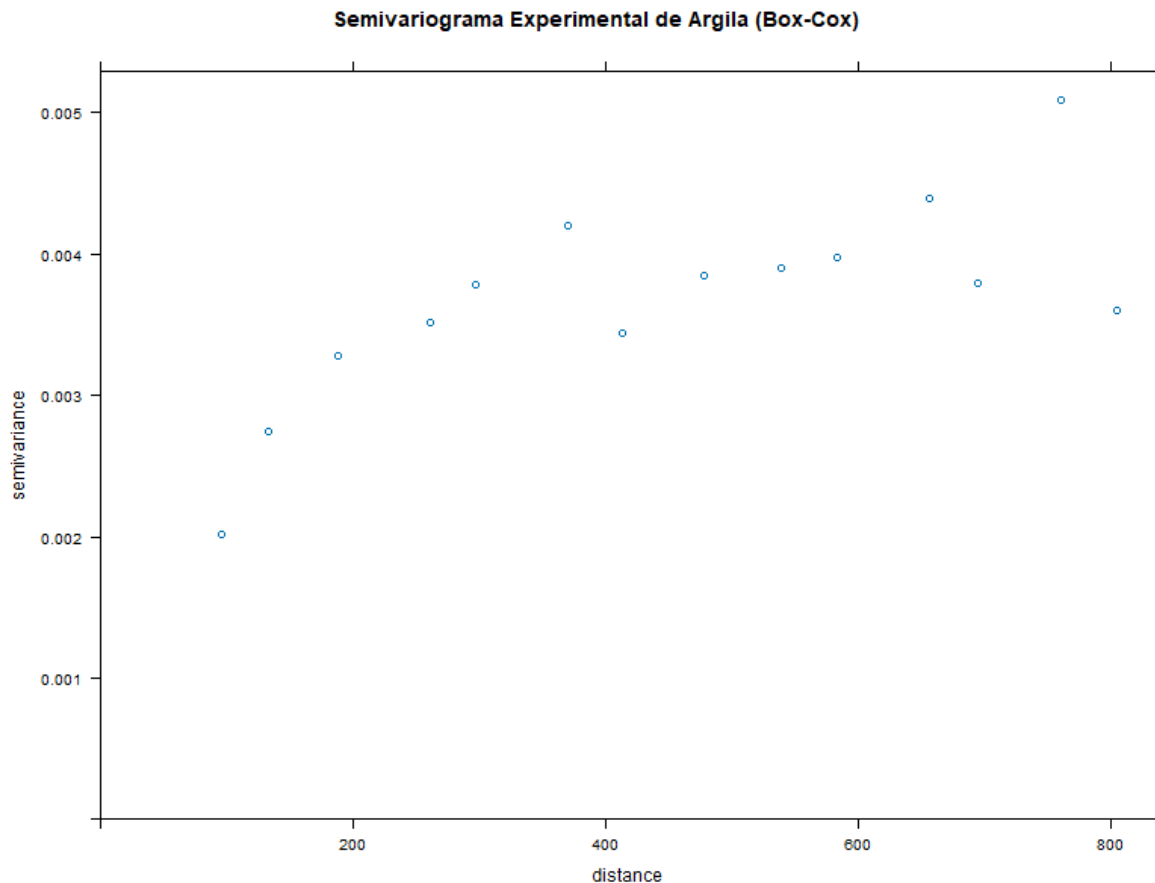
A equação básica utilizada para calcular a semivariância ( $\gamma(h)$ ) é:

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

onde:

- $h$ : distância entre pares de pontos;
- $N(h)$ : número de pares de pontos separados por  $h$ ;
- $Z(x_i)$ : valor da variável na posição  $x_i$ .





*Figura 3: Semivariograma Experimental*

## 5.2. Ajuste do Modelo Teórico

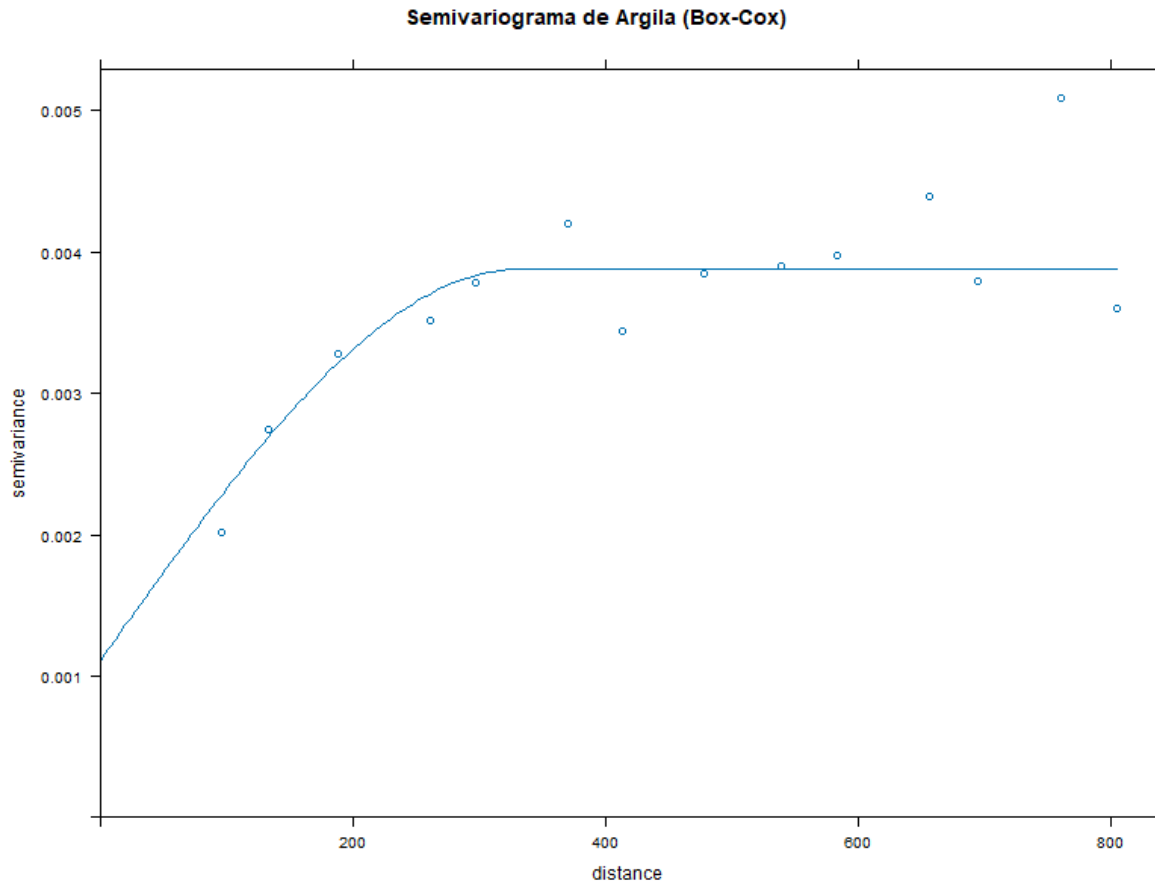
Após a obtenção do semivariograma experimental, foi ajustado um **modelo teórico** (Figura 4) para descrever a estrutura de dependência espacial. O modelo selecionado foi o **modelo esférico**, com os seguintes parâmetros ajustados:

- **Nugget (C0\_00):** Representa a variabilidade não explicada, ou erro de medição.
- **Sill (C0\_00 + C1\_11):** Variância total atingida quando a dependência espacial deixa de existir.
- **Range (a):** Distância na qual a semivariância atinge o sill, indicando o limite da dependência espacial.

A escolha do modelo esférico baseou-se na observação visual da forma do semivariograma experimental, onde o aumento da semivariância é mais acentuado inicialmente e estabiliza após o range.

Os parâmetros ajustados foram:

- **Nugget:** 0.1
- **Sill:** 0.6
- **Range:** 500 metros



*Figura 4: Semivariograma Ajustado*

### 5.3. Resultados

A análise variográfica revelou que os dados apresentam uma dependência espacial significativa até aproximadamente **500 metros**. Após essa distância, a semivariância estabiliza, indicando que os valores de argila transformada em pontos distantes mais de 500 metros são essencialmente independentes.

Este modelo serviu de base para as próximas etapas, particularmente a interpolação espacial por krigagem, que explora a dependência espacial para estimar valores em locais não amostrados.

## 6. Interpolação por Krigagem

Após a análise variográfica e o ajuste do modelo esférico, foi realizada a interpolação por krigagem para estimar os valores da variável transformada Argila\_boxcox em locais não amostrados. A krigagem é uma técnica geoestatística baseada na teoria das variáveis regionalizadas, que utiliza a dependência espacial identificada no semivariograma para gerar estimativas otimizadas.

### 6.1. Metodologia

#### 6.1.1. Configuração da Grade

Foi definida uma grade regular com resolução de **5 metros**, cobrindo toda a extensão do polígono limite da área estudada. A escolha dessa resolução visou um equilíbrio entre a qualidade das estimativas e a eficiência computacional.

A grade foi criada a partir da extensão do shapefile do imóvel, garantindo que todas as estimativas se limitassem à área de interesse.

#### 6.1.2. Modelo Utilizado

O modelo esférico ajustado na análise variográfica foi aplicado para realizar a interpolação, considerando os parâmetros:

- **Nugget:** 0.1
- **Sill:** 0.6
- **Range:** 500 metros.

#### 6.1.3. Processo de Krigagem

A interpolação foi realizada com o método de krigagem ordinária. Este método assume que a média local dos valores é desconhecida, mas constante dentro do alcance do modelo variográfico. A equação utilizada na krigagem ordinária é:

$$\hat{Z}(x_0) = \sum_{i=1}^N \lambda_i Z(x_i)$$

#### 6.1.4. Corte pelo Polígono Limite

O raster resultante da krigagem foi recortado para se ajustar ao polígono do imóvel, restringindo as estimativas à área de estudo.

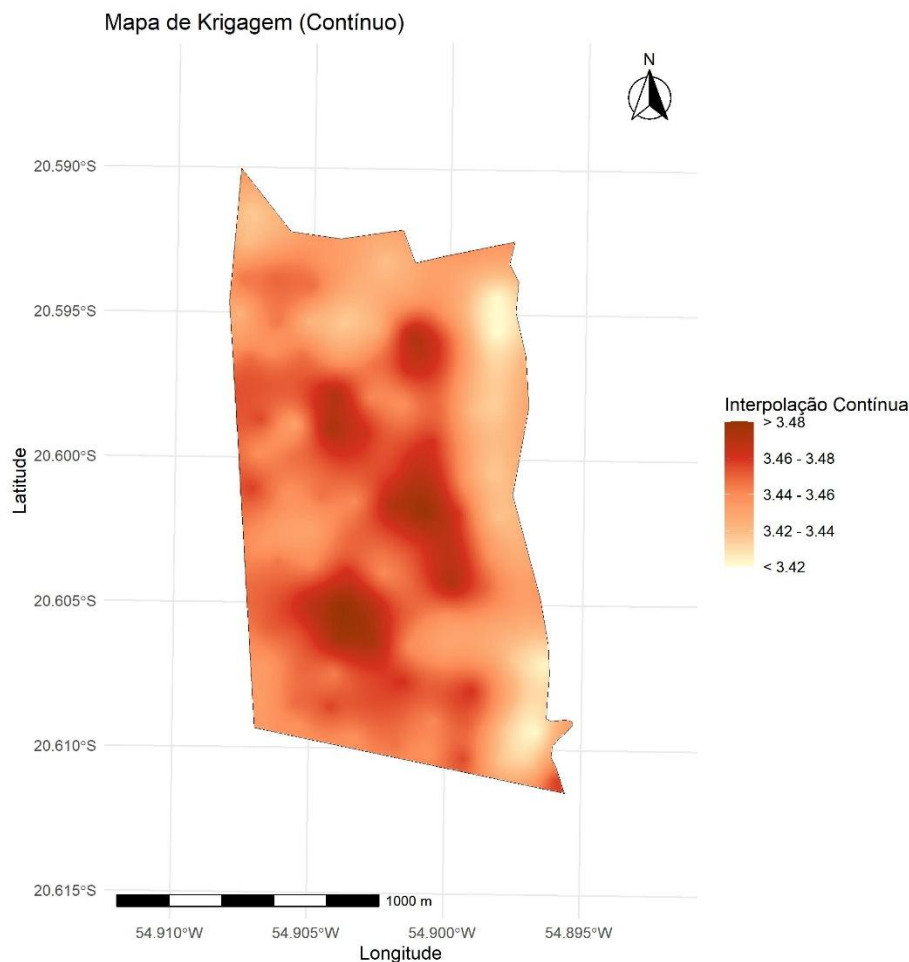
## 6.2. Resultados

A interpolação resultou em uma superfície contínua representando os valores estimados de argila transformada. O mapa gerado foi posteriormente classificado em intervalos, com a finalidade de facilitar a interpretação prática.

## 6.2.1. Produtos Gerados

### 6.2.1.1. Mapa de Krigagem Contínua (Figura 5)

Este mapa apresenta os valores interpolados como uma superfície contínua, evidenciando gradientes e padrões espaciais da distribuição de argila transformada.



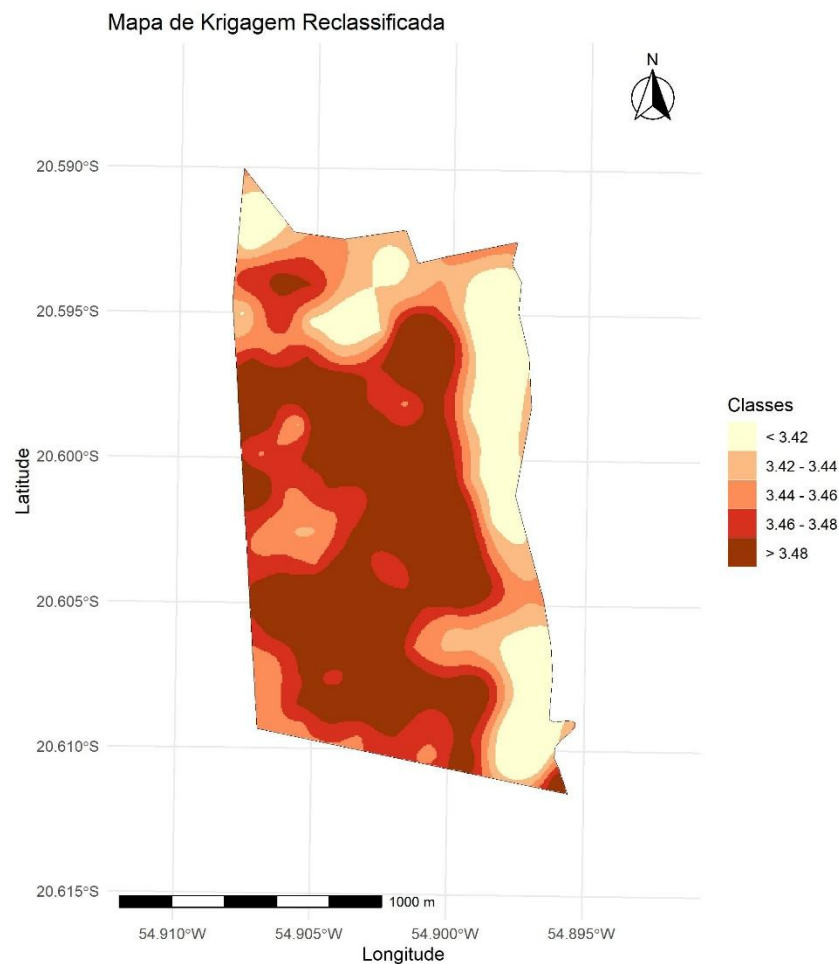
**Figura 5: Mapa de Krigagem Contínua.**

### 6.2.1.2. Mapa de Krigagem Reclassificada (Figura 6)

Os valores interpolados foram categorizados em cinco classes, com os seguintes intervalos:

- < 3.42
- 3.42 - 3.44
- 3.44 - 3.46
- 3.46 - 3.48
- > 3.48

Essa reclassificação auxilia na identificação de regiões com características homogêneas para aplicações práticas, como planejamento de manejo ou uso do solo.



**Figura 6: Mapa de Krigagem Reclassificada.**

### 6.3. Considerações

Os resultados da krigagem foram consistentes com as características observadas nos dados originais e na análise variográfica, validando o uso do modelo esférico e da técnica de krigagem

ordinária. A interpolação forneceu uma visão detalhada da variabilidade espacial da argila transformada, contribuindo para decisões informadas na gestão da área estudada.

## 7. Mapa de Variabilidade Espacial

Após a interpolação por krigagem, foi gerado o Mapa de Variabilidade Espacial utilizando o Coeficiente de Variação (CV), que é uma métrica relativa que relaciona o desvio padrão à média dos valores interpolados. Este indicador permite identificar áreas com maior ou menor dispersão nos valores estimados, auxiliando na compreensão da heterogeneidade espacial dos dados.

### 7.1. Metodologia

#### 7.1.1. Cálculo do Coeficiente de Variação

O CV foi calculado diretamente sobre os valores estimados no raster gerado pela krigagem:

$$CV = \left( \frac{\text{Desvio Padrão}}{\text{Média}} \right) \times 100$$

Essa fórmula foi aplicada a cada célula do raster, considerando os valores da interpolação na região de interesse.

#### 7.1.2. Representação Espacial

O raster resultante foi recortado utilizando o polígono limite da área estudada, garantindo que apenas a região dentro do imóvel fosse analisada.

#### 7.1.3. Visualização e Escala de Cores

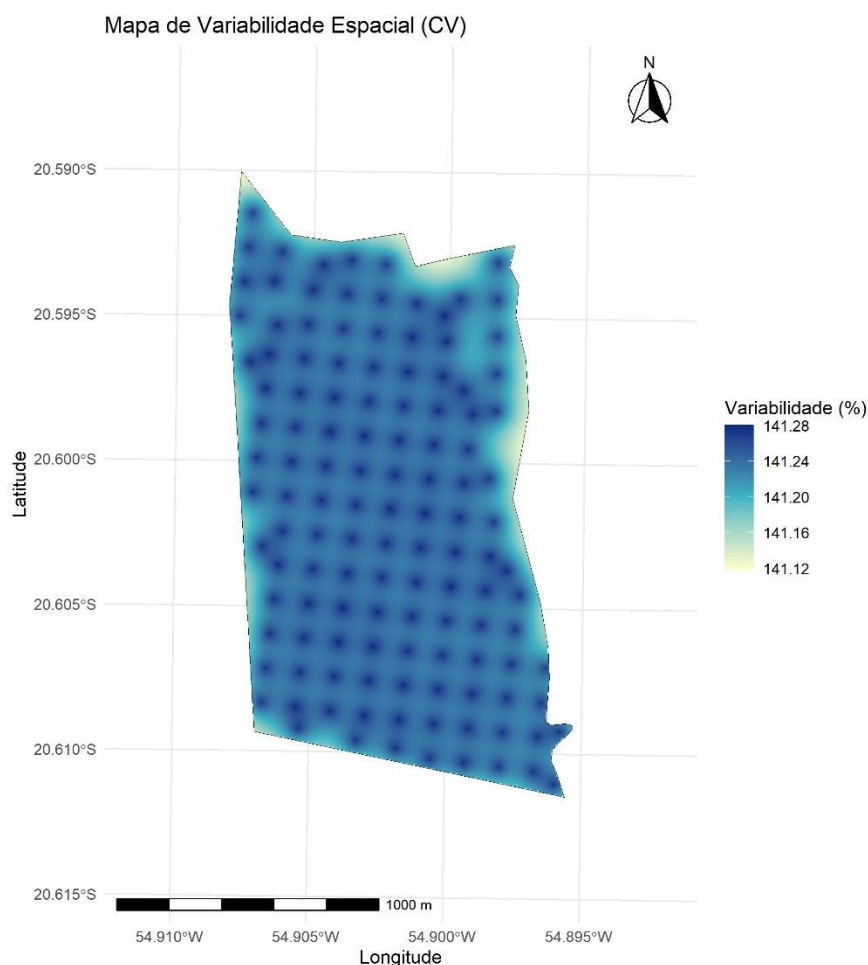
O mapa foi produzido com uma escala de cores contínua, indo do amarelo (baixa variabilidade) ao azul profundo (alta variabilidade). Isso facilita a identificação visual de áreas de interesse.

#### 7.1.4. Ajustes de Borda

Para assegurar uma apresentação clara e consistente, os limites do polígono foram ampliados em 20% com relação à extensão original, adicionando uma margem para melhor centralização no mapa.

## 7.2. Resultados

O Mapa de Variabilidade Espacial (CV) (Figura 7) destaca áreas com diferentes graus de dispersão nos valores estimados, fornecendo informações importantes para a análise e interpretação da variabilidade da argila transformada.



*Figura 7: Mapa de Variabilidade Espacial (CV).*

## 7.3. Considerações

O uso do coeficiente de variação como métrica relativa foi essencial para evitar a influência de diferenças absolutas nos valores estimados. Essa abordagem permitiu um entendimento mais refinado da variabilidade espacial, complementando os resultados obtidos com a krigagem.

## 8. Discussões



O uso do R para análises geoespaciais mostrou-se eficiente e flexível ao longo do presente estudo, permitindo integrar todas as etapas da análise, desde a importação de dados até a geração de mapas detalhados de variabilidade espacial e interpolação por krigagem. A aplicação de bibliotecas específicas como sf, raster, gstat, e ggplot2 evidenciou a robustez da linguagem para manipular dados espaciais em diferentes formatos, realizar cálculos avançados e produzir visualizações de alta qualidade. A flexibilidade do R em personalizar métodos e resultados destacou-se como um diferencial em comparação a softwares SIG, como ArcGIS e QGIS, que, embora tenham interfaces mais intuitivas, podem ser limitados em relação à customização e automação de processos (Bivand et al., 2013).

Os resultados alcançados confirmaram a adequação das técnicas utilizadas, especialmente no que se refere à representação da variabilidade espacial do teor de argila. A transformação Box-Cox foi eficaz para estabilizar a variância dos dados e aproximá-los de uma distribuição normal, mesmo que a normalidade completa não tenha sido alcançada. O ajuste do semivariograma e a interpolação por krigagem possibilitaram identificar padrões espaciais importantes, demonstrando a confiabilidade das análises geoespaciais conduzidas. Além disso, os mapas gerados permitiram uma interpretação clara e objetiva dos resultados, corroborando a relevância da abordagem estatística empregada (Webster & Oliver, 2007).

Apesar das vantagens evidenciadas, algumas limitações foram observadas durante o uso do R. A curva de aprendizado associada à linguagem pode ser desafiadora para usuários iniciantes, especialmente devido à necessidade de compreender conceitos estatísticos avançados e a sintaxe específica dos pacotes. Ademais, a manipulação de grandes volumes de dados geoespaciais pode demandar recursos computacionais elevados, o que foi mitigado neste estudo pela escolha de resoluções e parâmetros adequados. Ainda assim, o desempenho geral do R foi consistente e está alinhado com as práticas recomendadas na literatura.

A experiência de utilizar o R para análises geoespaciais reafirma sua aplicabilidade em estudos ambientais. As etapas realizadas, como o ajuste de modelos variográficos e a interpolação espacial, não apenas forneceram resultados confiáveis, mas também permitiram insights valiosos sobre a distribuição espacial dos atributos do solo na área estudada. Essa abordagem destaca o potencial do R como ferramenta central para estudos geográficos, desde que os métodos sejam aplicados de maneira criteriosa e bem fundamentada (Pebesma, 2018).

## 9. Conclusão





O estudo demonstrou a eficácia do R como uma ferramenta robusta e versátil para análises geoespaciais, integrando diferentes métodos estatísticos e de geoprocessamento em um único ambiente de programação. Os resultados obtidos, como os mapas de variabilidade e interpolação por krigagem, mostraram-se coerentes e interpretáveis, atendendo plenamente aos objetivos propostos. A análise variográfica, fundamentada em modelos teóricos ajustados, confirmou a confiabilidade das técnicas empregadas e ressaltou a importância de parâmetros bem definidos para a qualidade dos resultados.

Embora a utilização do R apresente desafios, como a necessidade de maior conhecimento técnico e recursos computacionais adequados, suas vantagens superam as limitações, especialmente em contextos onde a customização e a reprodutibilidade das análises são indispensáveis. A comparação com ferramentas proprietárias evidenciou que o R pode ser não apenas uma alternativa viável, mas também uma solução preferencial para análises geoespaciais avançadas.

Por fim, o presente estudo contribui para ampliar a adoção do R em estudos ambientais e reforça sua aplicabilidade para análises complexas. A experiência adquirida neste trabalho poderá ser replicada em estudos futuros, incluindo aplicações em diferentes variáveis ambientais e escalas espaciais. A integração de métodos estatísticos e geoespaciais no R, como a realizada aqui, destaca seu papel central no avanço das ciências ambientais e geográficas.

## 10. Disponibilidade de Dados e Código

Todos os arquivos utilizados neste estudo, incluindo o código completo, resultados, textos citados, camadas geoespaciais e este relatório, estão disponíveis em um repositório público no GitHub. O repositório pode ser acessado através do link: <https://github.com/samuel-c-santos/ufra-analise-solos-argila>. A disponibilização do repositório visa promover a reprodutibilidade dos resultados apresentados e fornecer recursos úteis para outros pesquisadores e profissionais interessados em análises semelhantes.

## 11. Referências

BIVAND, R. S.; PEBESMA, E.; GOMEZ-RUBIO, V. **Applied Spatial Data Analysis with R**. 2. ed. Springer, 2013.

ISAAKS, E. H.; SRIVASTAVA, R. M. **An Introduction to Applied Geostatistics**. New York: Oxford University Press, 1989.

MCBRATNEY, A. B.; WEBSTER, R. **Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates**. Journal of Soil Science, v. 37, n. 4, p. 617-639, 1986. Disponível em: <https://doi.org/10.1111/j.1365-2389.1986.tb00392.x>, acesso em 31/12/2024.

OSBORNE, J. W.; CHRISTIANSEN, W. R. I.; GUNTER, J. S. **The power of outliers (and why researchers should ALWAYS check for them)**. *Practical Assessment, Research & Evaluation*, v. 9, n. 6, p. 1-9, 2004. Disponível em: <https://doi.org/10.7275/qf69-7k43>, acesso em 31/12/2024.

PEBESMA, E. **Simple Features for R: Standardized Support for Spatial Vector Data**. The R Journal, v. 10, n. 1, p. 439–446, 2018.

WEBSTER, R.; OLIVER, M. A. **Geostatistics for environmental scientists**. 2. ed. Chichester: John Wiley & Sons, 2007. Disponível em: <https://doi.org/10.1002/9780470517277>, acesso em 31/12/2024.

WEBSTER, R.; OLIVER, M. A. **Geostatistics for Environmental Scientists**. 2. ed. Wiley-Blackwell, 2007.