

UNIVERSIDADE DE SÃO PAULO  
ESCOLA POLITÉCNICA

RAFAELLA MURARO

**Aplicação de simulação sequencial gaussiana na caracterização da variabilidade espacial  
da concentração de chumbo no solo.**

São Paulo  
2022

**Aplicação de simulação sequencial gaussiana na caracterização da variabilidade espacial da concentração de chumbo no solo.**

Versão Original

Monografia apresentada à Escola Politécnica da Universidade de São Paulo como parte dos requisitos para a obtenção do título de Especialista em Gestão de Áreas Contaminadas, Desenvolvimento Urbano Sustentável e Revitalização de Brownfields.

Orientador: Prof. Dr. Marcelo Monteiro da Rocha

São Paulo

2022

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

#### Catálogo na Publicação



## AGRADECIMENTOS

Gostaria de agradecer a todos envolvidos na elaboração desse trabalho e a todos que me apoiaram ao longo do curso.

Ao Professor Marcelo Monteiro da Rocha pela orientação, ensinamentos e atenção durante a elaboração desse trabalho.

Ao Antonio Tadashi do laboratório de informática do Instituto de Geociências pelo apoio.

Aos colegas da ERM pela compreensão e incentivo.

À minha família, Giuliano, Dirce e Clorinda, por sempre me incentivarem a estudar e por terem tornado possível a conclusão desse curso.

## RESUMO

MURARO, Rafaella. Aplicação de simulação sequencial gaussiana na caracterização da variabilidade espacial da concentração de chumbo no solo. São Paulo, 2022. 35 f. Monografia (MBA em Gestão de Áreas Contaminadas, Desenvolvimento Urbano Sustentável e Revitalização de Brownfields) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2022.

A caracterização da distribuição espacial da concentração de contaminantes em subsuperfície é uma etapa crucial no processo de gerenciamento de áreas contaminadas. É a partir da modelagem espacial do contaminante que serão definidas áreas com existência de risco de superação de valores de referência e a consequente necessidade de remediação do terreno, além de quais seriam as melhores técnicas a serem aplicadas de acordo com as características do local e das substâncias químicas de interesse.

Dentre as técnicas geoestatísticas, a Simulação Sequencial Gaussiana apresenta a vantagem de possuir precisão global, sendo capaz de reproduzir a heterogeneidade do fenômeno, honrando os dados amostrais e a correlação espacial estabelecida pela função variograma.

Neste contexto, o presente trabalho teve como objetivo a aplicação da Simulação Sequencial Gaussiana na caracterização da distribuição da concentração de chumbo no solo da região de Jura, na Suíça, em uma área de aproximadamente 14.5 km<sup>2</sup>. A partir de 100 realizações obtidas, foram gerados mapas de melhor e pior cenário, médias das concentrações simuladas e mapas de probabilidade. Dessa forma foi possível realizar a análise probabilística do risco de extrapolação dos valores de referência para chumbo no solo, considerando uma abordagem conservadora e outra não conservadora.

## ABSTRACT

MURARO, Rafaella. Spatial variability characterization of lead concentration in soil through the application of the Sequential Gaussian Simulation method. São Paulo, 2022. 35 f. Monography (MBA em Gestão de Áreas Contaminadas, Desenvolvimento Urbano Sustentável e Revitalização de Brownfields) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2022.

The spatial distribution characterization of contaminants concentrations in the subsurface is a crucial step in the process of managing contaminated areas. The modeling of spatial distribution of contaminant makes it possible to classify areas with risk of exceedances of guiding values, and its consequent need for land remediation, in addition to guide the selection of the best techniques to be applied according to the characteristics of the land and the chemical substances of interest.

Among the geostatistical techniques, the Sequential Gaussian Simulation has the advantage of having global precision, being able to reproduce the heterogeneity of the phenomenon, honoring the sample data and the spatial correlation established by the variogram function.

In this context, the present work aimed to characterize the distribution of lead concentration in the soil by applying the Sequential Gaussian Simulation method in Jura region, Switzerland, in an area of approximately 14,5 km<sup>2</sup>. From 100 realizations obtained, maps of best and worst scenario, averages of simulated concentrations and probability maps were generated. In this way, it was possible to conduct a probabilistic analysis of the risk of extrapolation of reference values for lead in the soil, considering a conservative and a non-conservative approach.

## ÍNDICE DE FIGURAS

|                                                                                                                                                                                                                                                  |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figura 1- Duas direções de cálculo do variograma experimental e suas respectivas tolerâncias. O triângulo preenchido em rosa é a direção de 45° com tolerância de 22,5° e o triângulo em azul é a direção de 90° também com tolerância de 22,5°. | 6  |
| Figura 2 - Variogramas experimentais para as direções 0° e 90°                                                                                                                                                                                   | 7  |
| Figura 3 - Variograma com patamar e demais propriedades                                                                                                                                                                                          | 8  |
| Figura 4 – Tipos de anisotropias existentes nos variogramas                                                                                                                                                                                      | 9  |
| Figura 5 - Transformação Gaussiana por meio da função de distribuição acumulada                                                                                                                                                                  | 13 |
| Figura 6 – Sequência de etapas metodológicas previstas para o trabalho.                                                                                                                                                                          | 17 |
| Figura 7 – Histograma da variável chumbo.                                                                                                                                                                                                        | 18 |
| Figura 8 – Mapa de pontos dos dados amostrais.                                                                                                                                                                                                   | 19 |
| Figura 9 – Histograma da variável chumbo transformada por <i>nscore</i> .                                                                                                                                                                        | 20 |
| Figura 10 - Variograma experimental e modelo de variograma ajustado para a direção 0°.                                                                                                                                                           | 21 |
| Figura 11 - Variograma experimental e modelo de variograma ajustado para a direção 90°.                                                                                                                                                          | 22 |
| Figura 12: Diagrama de dispersão entre os dados amostrais e estimados transformados por <i>normal score</i> .                                                                                                                                    | 23 |
| Figura 13: Realização n° 78 para a variável chumbo, resultante do método de SGS.                                                                                                                                                                 | 24 |
| Figura 14: Mapa de pior cenário para a variável chumbo                                                                                                                                                                                           | 25 |
| Figura 15: Histograma dos dados simulados para o cenário de maiores concentrações de chumbo.                                                                                                                                                     | 25 |
| Figura 16: Gráfico de distribuição de frequências acumuladas para os valores máximos simulados para cada nó.                                                                                                                                     | 26 |
| Figura 17: Mapa de melhor cenário para a variável chumbo                                                                                                                                                                                         | 27 |
| Figura 18 - Histograma dos dados simulados para o cenário de menores concentrações de chumbo.                                                                                                                                                    | 27 |
| Figura 19: Gráfico de distribuição de frequências acumuladas para os valores mínimos simulados para cada nó.                                                                                                                                     | 28 |
| Figura 20: Mapa de médias obtido a partir do conjunto de simulações.                                                                                                                                                                             | 29 |
| Figura 21: Histograma dos dados médios obtidos a partir do conjunto de realizações para a variável chumbo                                                                                                                                        | 29 |
| Figura 22: Gráfico de distribuição de frequências acumuladas para as médias de valores simulados para cada nó.                                                                                                                                   | 30 |

|                                                                                                    |    |
|----------------------------------------------------------------------------------------------------|----|
| Figura 23: Mapa de probabilidade para o valor de 50 mg/kg.....                                     | 31 |
| Figura 24: Mapa de probabilidade para o valor de 72 mg/kg.....                                     | 32 |
| Figura 25: Mapa de probabilidade para o valor de 150 mg/kg.....                                    | 33 |
| Figura 26 – Mapa de coeficientes de variação .....                                                 | 33 |
| Figura 27 – Mapa de risco de extrapolação dos valores orientadores – Cenário conservador .         | 35 |
| Figura 28 – Mapa de risco de extrapolação dos valores orientadores – Cenário não conservador ..... | 35 |

## ÍNDICE DE TABELAS

|                                                                                                                          |    |
|--------------------------------------------------------------------------------------------------------------------------|----|
| Tabela 1 – Estatística Descritiva dos dados de Chumbo (mg/kg). ....                                                      | 19 |
| Tabela 2 - Estatística Descritiva dos dados de Chumbo transformada por <i>normal score</i> .....                         | 20 |
| Tabela 3 – Parâmetros utilizados da análise exploratória. ....                                                           | 21 |
| Tabela 4 – Estruturas dos modelos de variogramas experimentais .....                                                     | 22 |
| Tabela 5 – Parâmetros de busca aplicados na SGS. ....                                                                    | 23 |
| Tabela 6 – Estatística descritiva do cenário de maiores concentrações para chumbo. ....                                  | 25 |
| Tabela 7 - Percentis correspondentes aos valores de referência para o pior cenário simulado                              | 26 |
| Tabela 8 – Estatística descritiva do cenário de menores concentrações para chumbo. ....                                  | 27 |
| Tabela 9 - Frequências acumuladas correspondentes aos valores de referência para o melhor cenário simulado. ....         | 28 |
| Tabela 10 – Estatística descritiva da média dos dados simulados. ....                                                    | 29 |
| Tabela 11 – Percentis correspondentes aos valores de referência. ....                                                    | 30 |
| Tabela 12 – Porcentagens da área total com probabilidade de extrapolação do limite de 50 mg/kg.....                      | 31 |
| Tabela 13 – Porcentagens da área total com probabilidade de extrapolação do limite de 72 mg/kg.....                      | 32 |
| Tabela 14 – Porcentagens da área total com risco de extrapolação dos limites orientadores. .                             | 35 |
| Tabela 15 – Porcentagens da área total com risco de extrapolação dos limites orientadores – Cenário não conservador..... | 36 |

## SUMÁRIO

|        |                                                                                      |    |
|--------|--------------------------------------------------------------------------------------|----|
| 1.     | INTRODUÇÃO .....                                                                     | 1  |
| 2.     | OBJETIVOS .....                                                                      | 2  |
| 3.     | JUSTIFICATIVA .....                                                                  | 3  |
| 4.     | REVISÃO BIBLIOGRÁFICA PRELIMINAR .....                                               | 4  |
| 4.1.   | Análise geoestatística .....                                                         | 4  |
| 4.1.1. | Variograma.....                                                                      | 5  |
| 4.1.2  | Anisotropias.....                                                                    | 8  |
| 4.2.   | Validação cruzada.....                                                               | 9  |
| 4.3.   | Métodos geoestatísticos .....                                                        | 9  |
| 4.3.1  | Krigagem .....                                                                       | 9  |
| 4.3.2  | Simulação Estocástica .....                                                          | 11 |
| 4.4.   | Trabalhos prévios – Aplicação de métodos geoestatísticos em estudos ambientais ..... | 14 |
| 5.     | MATERIAIS E MÉTODOS .....                                                            | 16 |
| 5.1.   | Base de dados .....                                                                  | 16 |
| 5.2.   | Métodos .....                                                                        | 16 |
| 6.     | RESULTADOS E DISCUSSÃO .....                                                         | 18 |
| 6.1.   | Análise Exploratória dos dados .....                                                 | 18 |
| 6.1.1  | Análise Estatística e Geoestatística dos dados .....                                 | 18 |
| 6.2.   | Validação Cruzada .....                                                              | 22 |
| 6.3.   | Simulação Sequencial Gaussiana .....                                                 | 23 |
| 6.3.1. | Simulações de melhor e pior cenário .....                                            | 24 |
| 6.3.2. | Média do conjunto de realizações .....                                               | 28 |
| 6.3.3. | Mapas de probabilidade de extrapolação dos valores orientadores .....                | 30 |
| 6.3.4. | Mapas de risco .....                                                                 | 34 |
| 7.     | CONSIDERAÇÕES FINAIS .....                                                           | 37 |
| 8.     | REFERÊNCIAS BIBLIOGRÁFICAS .....                                                     | 38 |

## 1. INTRODUÇÃO

A caracterização da distribuição espacial da concentração de contaminantes em subsuperfície é uma etapa crucial no processo de gerenciamento de áreas contaminadas. É a partir da modelagem espacial do contaminante que serão definidas áreas com existência de risco à saúde humana e a consequente necessidade de remediação do terreno, além de quais seriam as melhores técnicas a serem aplicadas de acordo com as características do local e das substâncias químicas de interesse.

Neste contexto, os métodos geoestatísticos apresentam-se como ferramentas eficazes e adequadas para a estimativa da variável de interesse em pontos não amostrados do espaço, quando comparados aos métodos estatísticos clássicos, uma vez que a análise geoestatística possibilita a caracterização da variabilidade espacial, fator intrínseco à maioria dos fenômenos naturais (Isaaks e Srivastava, 1989). A partir do conceito de variáveis regionalizadas, cunhado por Matheron (1965), a correlação espacial entre os dados de uma variável é apresentada por meio de uma função denominada de variograma.

A krigagem é o procedimento geoestatístico utilizado com maior frequência como ferramenta no gerenciamento de áreas contaminadas. No entanto, um problema relacionado à essa técnica é o efeito de suavização das estimativas, o qual provoca a diminuição da variabilidade espacial do fenômeno tratado, sendo a simulação estocástica a solução adotada para resolver este impasse (Yamamoto & Landim, 2013). A simulação gera vários cenários – diferentes entre si – e a partir deles pode-se quantificar a incerteza espacial, possibilitando a tomada de decisões com base na avaliação do risco associado, o que não é possível a partir apenas da estimativa.

Nesse contexto, o presente trabalho possui como objetivo a aplicação da Simulação Sequencial Gaussiana na caracterização da distribuição da concentração de chumbo no solo da região de Jura, na Suíça, em uma área de aproximadamente 14,5 km<sup>2</sup>. O banco de dados, composto por 259 pontos de amostragem, é de domínio público e foi obtido pelo Instituto Federal de Tecnologia, em Lausanne. A partir da simulação de 100 diferentes cenários equiprováveis serão definidos o pior e o melhor cenário da distribuição espacial da contaminação gerados pela combinação das simulações calculadas, os quais serão analisados a partir de mapas de probabilidades de extrapolação de limites críticos.

## **2. OBJETIVOS**

O objetivo principal desse trabalho é a aplicação da técnica de Simulação Sequencial Gaussiana (SGS) na caracterização da distribuição espacial da contaminação de chumbo no solo, a partir do banco de dados de Jura. A partir das 100 realizações obtidas, serão gerados mapas de melhor e pior cenário, mapa de médias das concentrações simuladas e mapas de probabilidade. Dessa forma pretende-se realizar a análise probabilística do risco de extrapolação dos valores de referência para chumbo.

### **3. JUSTIFICATIVA**

A relevância da aplicação de técnicas geoestatísticas no gerenciamento de áreas contaminadas justifica-se pela sua característica de representar a variabilidade do fenômeno a partir da correlação espacial entre os dados. Dentre as técnicas geoestatísticas, a Simulação Sequencial Gaussiana apresenta a vantagem de possuir precisão global, sendo capaz de reproduzir a heterogeneidade do fenômeno, honrando os dados amostrais e a correlação espacial estabelecida pela função variograma.

Tal procedimento, ainda muito pouco difundido na área ambiental, possibilita a criação de modelos mais representativos da realidade da variável e suas respectivas incertezas, além de auxiliar a tomada de decisões relacionadas ao gerenciamento do risco.

## 4. REVISÃO BIBLIOGRÁFICA

### 4.1. Análise geostatística

Conforme destacado por Rocha et al. (2016), o conceito de variável regionalizada foi introduzido por Matheron (1965) para designar uma função que possui um valor específico associado a cada local do espaço. Trata-se, portanto, de uma variável aleatória com posição definida no espaço, que pode ser escrita como  $Z(x)$ , onde  $Z$  é o valor que a variável em estudo assume no ponto  $x$ , representado por coordenadas  $E$ ,  $N$  e cota.

Como pode ser observado em Yamamoto (2001), as variáveis regionalizadas possuem características que as diferenciam da variável aleatória, usualmente empregada na estatística clássica, sendo elas: localização, suporte, continuidade e anisotropia.

- Localização: Uma amostra precisa apresentar localização no espaço ( $x, y$  e  $z$ );
- Suporte: identifica se a amostra foi coletada em um ponto, área ou volume e sua orientação no espaço;
- Continuidade: independente se a continuidade é grande ou pequena, na avaliação de fenômenos naturais, sempre haverá alguma correlação entre pontos conhecidos no domínio amostral;
- Anisotropia: Define-se anisotropia como sendo o comportamento distinto de uma variável regionalizada para diferentes direções, podendo apresentar variações suaves em uma direção e bruscas em outras.

Segundo Armstrong (1998), o estudo geostatístico de uma variável regionalizada deve se basear na hipótese intrínseca a qual, de acordo com Matheron (1965), garante que a variância espacial é em média a mesma em todo o domínio, na condição de que os pontos analisados estejam dispostos em uma mesma direção e separados por uma mesma distância. Esta hipótese também engloba o conceito de estacionariedade de segunda ordem, que considera a variância como sendo invariante à translação.

#### 4.1.1. Variograma

Conforme Armstrong (1998), o variograma é a ferramenta básica da geoestatística sendo utilizado na quantificação da correlação espacial entre as observações. De forma similar ao conceito de variância na estatística clássica, que consiste na média de diferenças quadráticas entre os valores amostrais e a média da distribuição, a função variograma  $2\gamma(h)$  estabelece a variância espacial entre pontos amostrais separados por uma distância  $h$ , utilizando como comparação valores vizinhos no lugar da média amostral (Yamamoto e Landim, 2013; Rocha et al., 2016), conforme apresentado a seguir:

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

Onde:

$n$ : número de pares de pontos amostrais separados pela distância  $h$ ;

$Z(x_i)$ : variável regionalizada no ponto  $x_i$

$Z(x_i + h)$ : é a mesma variável, porém a uma distância  $h$  do ponto  $Z(x_i)$

No entanto, nos estudos de variáveis regionalizadas, utiliza-se a função semivariograma que correspondente à metade da função variograma (Rocha et al, 2016), ou seja:

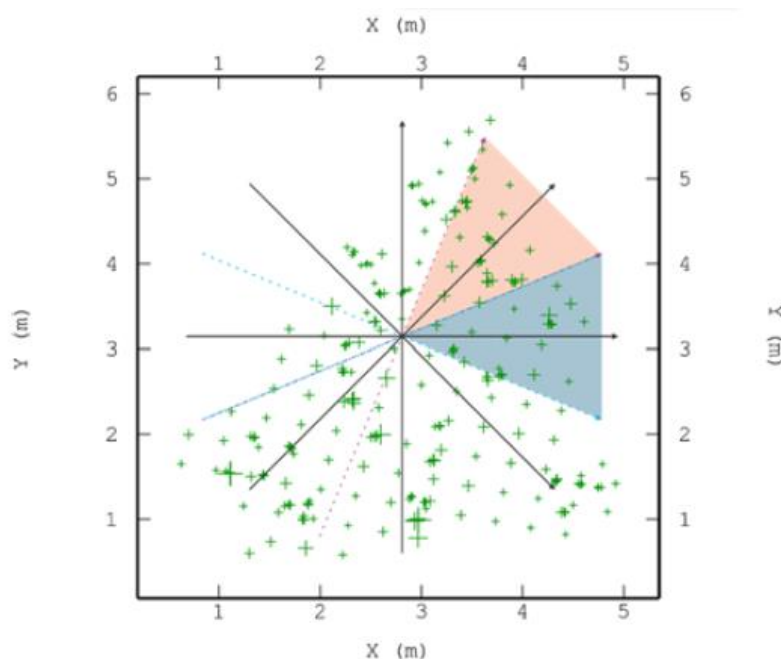
$$\gamma(h) = \frac{1}{2n} \sum_{i=1}^n [Z(x_i + h) - Z(x_i)]^2$$

Na literatura geoestatística tal função é usualmente denominada como variograma (Yamamoto e Landim, 2013), de forma que, no atual estudo será adotado o termo variograma para se referir à função semivariograma.

A análise geoestatística inicia-se com o cálculo do variograma experimental com o objetivo de reconhecer as anisotropias existentes e verificar o comportamento espacial do fenômeno. Para dados amostrais dispostos em malha irregular, como ocorre na maioria dos casos, este cálculo é realizado em quatro direções ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  e  $135^\circ$ ) na análise bidimensional e 5 direções azimutais ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ ,  $135^\circ$  e  $0^\circ$  com  $90^\circ$  de mergulho) na tridimensional. Adicionalmente, devem ser considerados os seguintes parâmetros: número de

passos, tamanho do passo, tolerância do passo e tolerância angular (Yamamoto e Landim, 2013; Rocha et al., 2016).

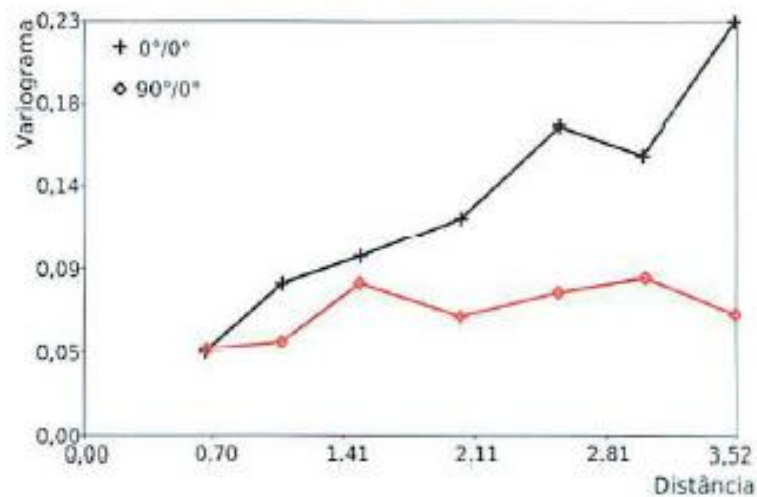
Figura 1- Duas direções de cálculo do variograma experimental e suas respectivas tolerâncias. O triângulo preenchido em rosa é a direção de 45° com tolerância de 22,5° e o triângulo em azul é a direção de 90° também com tolerância de 22,5°.



Fonte: Rocha et al. (2016)

Caso seja identificada a presença de anisotropia entre as direções do variograma experimental, é escolhida aquela de maior continuidade e então calcula-se o variograma experimental para duas ou três direções ortogonais (2D ou 3D, respectivamente) (Rocha et al, 2016). Na figura 2 é possível observar a presença de anisotropia visto que o variograma experimental mostra diferentes comportamentos nas diferentes direções de análise.

Figura 2 - Variogramas experimentais para as direções 0° e 90°



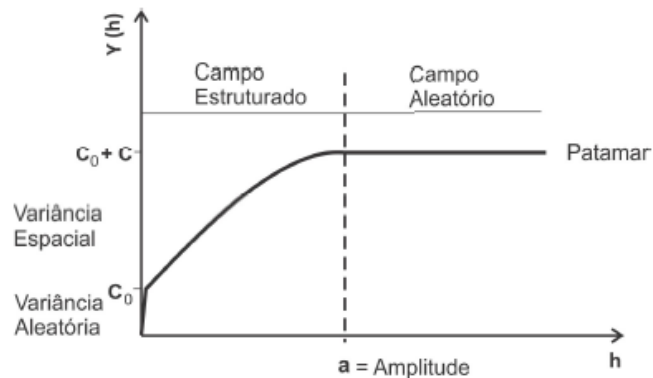
Fonte: Yamamoto e Landim (2013)

Após o cálculo dos variogramas experimentais, é realizado o ajuste do modelo teórico de variograma ao variograma experimental. O ajuste do modelo permite a obtenção da variância espacial para qualquer distância, visto que o variograma experimental fornece somente a variância de distâncias múltiplas.

O ajuste é realizado de forma interativa pelo usuário, que define qual função melhor se ajusta ao caso, além de quais os valores de parâmetros são os mais adequados para o modelo. As propriedades que caracterizam o variograma e deverão ser definidas são o efeito pepita, a variância espacial, a amplitude e o patamar (Yamamoto, 2001). Tais propriedades são definidas a seguir e ilustradas na Figura 3.

- Patamar ( $C_0 + C$ ): É a variância máxima na qual a variância espacial estabiliza-se. Em alguns casos a variância continua aumentando indefinidamente com a distância, nesses casos não há patamar. O patamar, também denominado de soleira, é dado pela notação  $C_0 + C$ , sendo  $C$  a variância espacial e  $C_0$  correspondente ao efeito pepita.
- Amplitude ( $a$ ): A distância na qual a função variograma atinge o patamar, também é chamada de alcance e representada pela notação “ $a$ ”.
- Efeito pepita ( $C_0$ ): O efeito pepita é o valor da variância espacial próximo à origem. O valor da variância espacial para distâncias nulas será igual a zero, porém na presença de efeito pepita, quando a distância tende a zero, o valor da variância será  $C_0$  devido ao efeito da variância aleatória. O efeito pepita pode ser resultante tanto da variabilidade natural do fenômeno espacial em estudo como da escala de amostragem.

Figura 3 - Variograma com patamar e demais propriedades



Fonte: Yamamoto (2001).

Embora existam diversos modelos teóricos de variograma, os modelos esférico, exponencial e gaussiano se ajustam a 90% dos variogramas calculados.

#### 4.1.2 Anisotropias

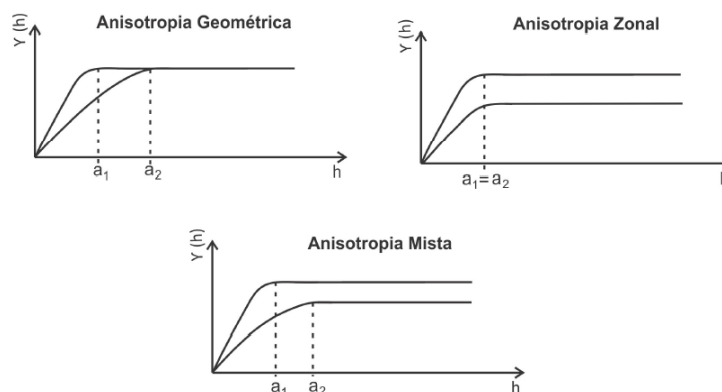
Conforme mencionado no **item 4.1.1**, com base no variograma obtido para diferentes direções, verifica-se a existência de anisotropia para a variável em questão, ou seja, diferentes comportamentos para diferentes direções. Quando não há distinção, a variável regionalizada possui comportamento isotrópico.

Existem três tipos distintos de anisotropia (Yamamoto e Landim, 2013), conforme exposto a seguir:

- Anisotropia zonal: apresenta patamares diferentes conforme a direção analisada, mas todos sob um mesmo alcance;
- Anisotropia mista: nesse caso, tanto a amplitude quanto o patamar variam conforme a direção;
- Anisotropia geométrica: caracteriza-se pela existência de um único patamar e amplitudes diferentes;

A Figura 4 ilustra variogramas com a presença das anisotropias mencionadas;

Figura 4 – Tipos de anisotropias existentes nos variogramas



Fonte: Yamamoto (2001).

## 4.2. Validação cruzada

A etapa de validação cruzada possui os objetivos de aferir o ajuste do modelo teórico de variograma realizado e, conseqüentemente, a qualidade do modelo interpretado, além de definir os melhores parâmetros de vizinhança a serem utilizados na estimativa por krigagem.

O procedimento ocorre por meio da krigagem pontual com a supressão do valor de um ponto amostral da base de dados, mantendo-se apenas sua posição. Em seguida, estima-se esse ponto por krigagem a partir das amostras situadas nos pontos vizinhos e o valor amostrado é devolvido àquele local. Este procedimento é repetido para todos os pontos amostrais, de forma a obter os erros da estimativa para cada ponto a partir da diferença entre o valor amostral e o valor estimado por krigagem (Armstrong, 1998).

Adicionalmente, a partir desse processo é possível obter a incerteza relacionada ao erro, calculada por meio da divisão do erro de estimativa pelo desvio padrão da interpolação. Essa medida é expressa em unidades de desvio padrão (Yamamoto e Landim, 2013).

## 4.3. Métodos geoestatísticos

### 4.3.1 Krigagem

De acordo com Armstrong (1998), a krigagem é um método linear de estimativa que fornece a melhor estimativa não tendenciosa e de mínima variância para o valor estimado da variável. Entre as diferentes técnicas existentes, a krigagem ordinária é a mais usual, no

entanto, no presente estudo, será utilizada a krigagem simples como uma das etapas da simulação sequencial gaussiana.

De acordo com Deutsch e Journel (1998), todas as versões da krigagem são baseadas no algoritmo básico de regressão linear e seu estimador correspondente. Sob as condições de estacionariedade de segunda ordem, com média ( $m$ ) e variância constantes no domínio, o estimador da krigagem simples é determinado a partir da seguinte expressão (conforme Yamamoto e Landim, 2013):

$$Z_{KS}^*(x_o) = m_0 + \sum_{i=1}^n \lambda_i [Z(x_i) - m_i]$$

Onde:

$Z_{KS}^*(x_o)$ : Estimador da krigagem simples

$m_0$ : é a média populacional

$m_i$ : é a média amostral assumida como estacionária

$Z(x_i)$ : Valor da variável nos pontos amostrados

$\lambda_i$ : ponderadores da krigagem simples

Para o cálculo dos ponderadores é definida uma nova função  $Y(x)$  que corresponde à diferença entre a função aleatória  $Z(x)$  e sua média (Armstrong, 1998; Yamamoto e Landim, 2013), ou seja:

$$Y(x) = Z(x) - E[Z(x)]$$

Onde:

$$E[Y(x)] = 0.$$

Com base na manipulação dessas equações a partir da covariância de  $Z(x)$  e  $Y(x)$  e visando a minimização da variância do erro, os ponderadores da krigagem simples são obtidos por meio de uma equação que pode ser expressa em termos matriciais e que não considera o multiplicador de Lagrange (Armstrong, 1998; Yamamoto e Landim, 2013), conforme:

$$\begin{bmatrix} C(x_1, x_1) & \cdots & C(x_1, x_n) \\ \vdots & \ddots & \vdots \\ C(x_n, x_1) & \cdots & C(x_n, x_n) \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} C(x_1, x_0) \\ \vdots \\ C(x_n, x_0) \end{bmatrix}$$

De acordo com Armstrong (1998), a krigagem simples raramente é utilizada em estudos de rotina, visto que seu emprego pressupõe o conhecimento da média populacional. Em casos como os de minas de grande porte situadas na África do Sul, a aplicação da krigagem simples ocorre com frequência, pois a média de cada área é conhecida devido à grande quantidade de amostras coletadas durante o longo tempo de operação dessas minas. Outro tipo de situação que permite o emprego dessa técnica é quando ocorre a transformação dos dados, que são manipulados de forma a obter valores de média populacional conhecidos, em geral iguais a zero.

#### 4.3.2 Simulação Estocástica

De acordo com Deutsch e Journel (1998), existem dois principais objetivos para a realização de simulações estocásticas. O primeiro objetivo consiste na criação de realizações que honrem o histograma, o variograma e os valores dos dados amostrais. Neste sentido, a simulação estocástica é considerada uma solução para o efeito de suavização causado pela krigagem, gerando realizações que representam adequadamente a variabilidade espacial do fenômeno tratado (precisão global), no entanto, isso ocorre em detrimento da precisão local.

O segundo objetivo da simulação estocástica é avaliar as incertezas espaciais por meio de diversas realizações. A partir do conjunto de realizações, é possível obter uma medida visual e quantitativa da incerteza espacial.

##### 4.3.2.1. Simulação Sequencial Gaussiana

Conforme Deutsch e Journel (1998), a Simulação Sequencial Gaussiana (SGS) consiste no algoritmo mais simples para geração de realizações a partir de simulação sequencial aplicada em funções aleatórias multigaussianas. Este procedimento resulta na geração de imagens equiprováveis, que representam a variabilidade espacial dos dados. No caso das simulações condicionais, as simulações são condicionadas pelos dados amostrais. No entanto, cabe ressaltar que na prática, em geral, os pontos amostrais não coincidem com os nós do grid utilizado, de forma que os valores amostrais serão atribuídos ao nó mais próximo.

O procedimento para realização da simulação condicional de uma variável contínua  $Z(x)$ , por meio da SGS, se dá por meio das seguintes etapas (Deutsch e Journel, 1998):

1. Determinar a função densidade acumulada da variável  $Z(x)$ ;
2. Usar a função de densidade acumulada da variável  $Z(x)$  para a transformação *normal score* da variável “ $z$ ” para a variável “ $y$ ”, que possui função densidade acumulada com distribuição normal;
3. Checar a normalidade bivariada da variável “ $y$ ” transformada. Se o modelo multigaussiano não puder ser estabelecido, considerar modelos alternativos, como a mistura de populações com distribuição gaussiana ou um algoritmo baseado em indicadores para a simulação estocástica;
4. Se for possível adotar o modelo da função aleatória multigaussiana para a variável “ $y$ ”:
  - Definir um caminho aleatório que visite cada nó do grid apenas uma vez. Para cada nó, estabelecer um número específico de dados condicionantes na vizinhança, considerando os dados amostrais e os valores previamente simulados;
  - Aplicar a krigagem simples com o modelo de semivariograma da variável, transformada por *normal score* (*n-score transformation*), para determinar os parâmetros (média e variância) da função de densidade acumulada da variável aleatória “ $Y(x)$ ” no ponto;
  - Extrair um valor simulado  $y^{(l)}(x)$  dessa função densidade acumulada condicional;
  - Adicionar o valor simulado  $y^{(l)}(x)$  ao conjunto de pontos de dados;
  - Repetir o procedimento para o próximo nó do grid, até que todos os nós tenham sido simulados;
5. Realizar a transformação inversa dos valores gaussianos simulados  $\{y^{(l)}(x), x \in A\}$  em valores simulados da variável original  $\{z^{(l)}(x) = \varphi^{-1}(y^{(l)}(x)), x \in A\}$ .

#### - Transformação por Normal Score (N-Score Transformation)

De acordo com Deutsch e Journel (1998), a primeira condição necessária para que a função aleatória estacionária,  $Y(x)$ , seja multigaussiana é que a sua função densidade acumulada possua distribuição normal (gaussiana), com média igual a 0 e variância igual a 1.

Distribuições não gaussianas podem ser transformadas em gaussianas, por meio da técnica *Normal Score* (*N-Score Transformation*). Nesta transformação, a função de

distribuição acumulada da variável a ser transformada e a função de distribuição acumulada gaussiana padronizada são utilizadas para realizar a transformação, conforme ilustrado na Figura 5.

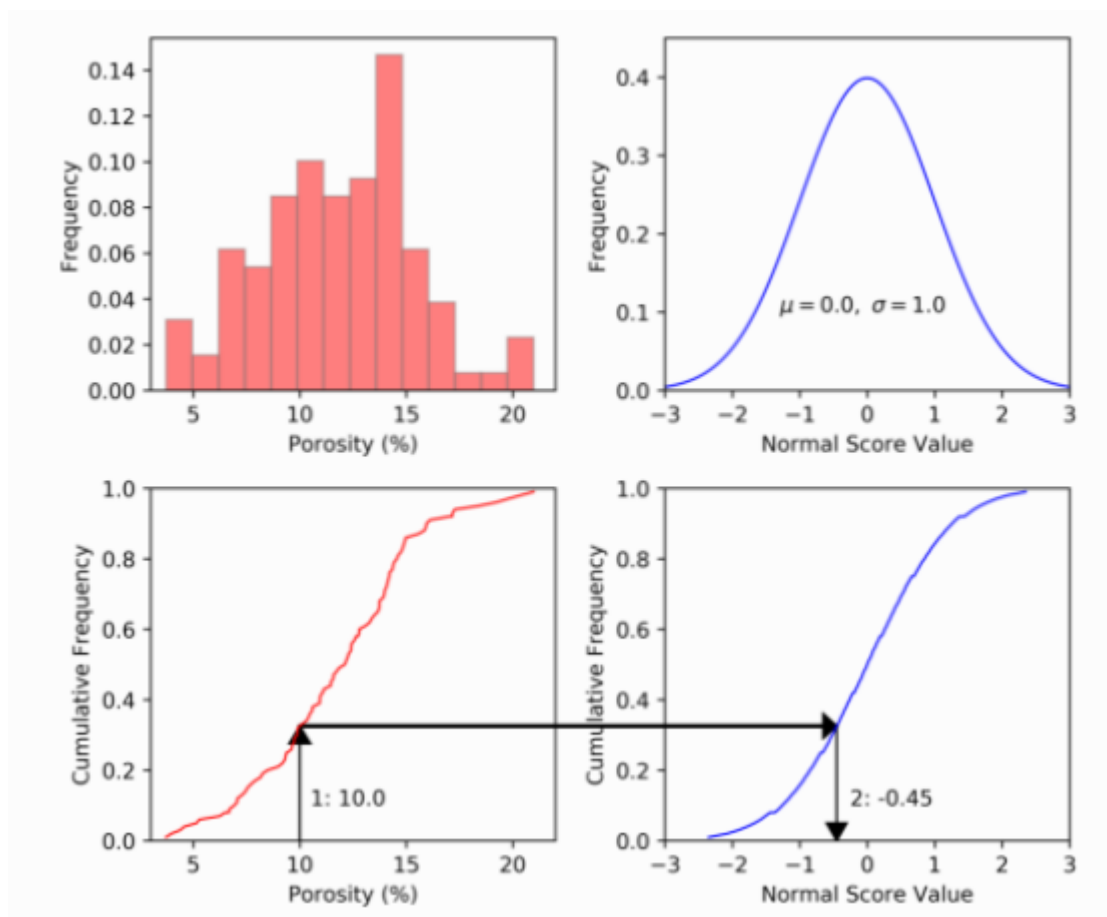
A transformação é dada pela transformação do quantil:

$$y = G^{-1}(F(z))$$

e pode ser inversamente transformada por:

$$z = F^{-1}(G(y))$$

Figura 5 - Transformação Gaussiana por meio da função de distribuição acumulada



Fonte: Pyrcz, e Deutsch (2018).

#### **4.4.Trabalhos prévios – Aplicação de métodos geoestatísticos em estudos ambientais**

Em estudos ambientais, é comum que as variáveis do fenômeno tratado possuam correlação espacial e temporal, como é o caso de amostras de concentrações de poluentes na subsuperfície, assunto tratado no presente trabalho. Desse modo, as técnicas geoestatísticas são as ferramentas mais apropriadas para elaboração de modelos de distribuição dessas variáveis e estimar os valores em localizações não amostradas, permitindo ainda a quantificação das incertezas de forma a possibilitar a avaliação de risco e o embasamento para tomadas de decisões.

Goovaerts (1997) mostra que o uso de mapas suavizados produzidos por métodos de interpolação como a krigagem é impróprio para determinadas aplicações, tais como a avaliação de fontes recuperáveis em depósitos minerais, modelamento de fluxos de fluidos em meios porosos ou delineamento de áreas contaminadas, pois ocorre a subestimativa dos valores extremos (*outliers*) e seus padrões de continuidade. Nesse contexto, a simulação estocástica tem sido preferida à krigagem em aplicações onde a variação espacial deve ser preservada (Queiroz, 2003). Apesar de não representar uma prática frequente no gerenciamento de áreas contaminadas, há um expressivo número de trabalhos que utilizam a técnica de simulação sequencial para avaliar fenômenos de contaminação.

Estudos a respeito da aplicação da abordagem probabilística para avaliação de contaminação em água subterrânea foram realizados por diversos autores. Istok e Rautmann (1996) discutem as vantagens da abordagem probabilística a partir da simulação estocástica, ao possibilitar a avaliação das incertezas e erros. Nesse trabalho foi utilizada a simulação condicional sequencial gaussiana para gerar mapas probabilísticos, a partir dos quais foram definidas a extensão e magnitude da contaminação de nitrato e Dacthal (herbicida) na zona saturada e não-saturada de uma área de aproximadamente 150 Km<sup>2</sup>, próxima a Ontario, Oregon.

Diversos estudos utilizam a simulação estocástica para caracterização da zona saturada, principalmente com simulação da porosidade com o objetivo de avaliar o transporte de fluxos no meio poroso, como é o caso dos trabalhos apresentados por Gomez-Hernandez et al (1997), Franssen e Gomez-Hernandez (2002) e Goovaerts (1999).

Goovaerts et al. (1997) utilizam indicadores geoestatísticos para avaliação do risco relacionado aos contaminantes no solo (cádmio, cobre e chumbo), na área de Jura, Suíça. Neste trabalho os autores utilizam a base de dados de amostragem que será utilizada no presente estudo, e que é amplamente utilizada em estudos geoestatísticos, a partir da aplicação de diversas ferramentas geoestatísticas entre as quais a krigagem ordinária, krigagem simples, krigagem indicativa, cokrigagem e simulações condicionais. O mesmo banco de dados foi previamente utilizado por Atteia et al. (1994) e Webster et al. (1994) na elaboração do mapa de estimativa das concentrações e na correlação entre as concentrações dos metais, respectivamente.

Ling e Chang (2000) avaliaram a distribuição espacial e magnitude das concentrações de zinco em áreas agrícolas de Taiwan, China, a partir da aplicação da krigagem ordinária e da simulação estocástica gaussiana. Os autores, assim como Goovaerts (1997), enfatizaram a vantagem da aplicação da simulação estocástica em estudos com grande variabilidade espacial, em comparação à krigagem, que causa grande efeito de suavização nos dados apresentados.

Broos et al. (1999) investigam os efeitos da variabilidade espacial de um solo contaminado com metais pesados, PAH e TPH no cálculo de custos para remediação. Os autores concluem sobre a importância de um modelo que leve em consideração a variabilidade espacial, como é o caso dos produtos gerados pela Simulação Sequencial Gaussiana, visto que de 8 a 13% das incertezas nos custos podem ser explicadas pela variabilidade espacial dos contaminantes e da geologia da área.

Entre estudos recentes que abordam a temática do gerenciamento de áreas contaminadas utilizando simulação, cita-se: Boente et al. (2019), Chen et al (2022), Lv e Liu (2019) e Metanhi (2019).

## 5. MATERIAIS E MÉTODOS

### 5.1. Base de dados

A técnica de Simulação Sequencial Gaussiana será aplicada a partir de dados de análises químicas em amostras de solo, na região de Jura, Suíça, em uma área de aproximadamente 14.5 km<sup>2</sup>. Os dados são de domínio público e foram obtidos pelo Instituto Federal de Tecnologia da Suíça, em Lausanne. O banco de dados original é constituído por 359 pontos amostrais e apresenta informações de análises químicas de metais pesados, entre os quais foi selecionado o chumbo como a variável de interesse para realizar o procedimento. Dentre o total de amostras, serão utilizadas as 259 amostras pertencentes ao grupo denominado como “grupo de predição”, sendo que as outras 100 pertencem ao “grupo de validação” e não serão consideradas no presente trabalho.

A partir da simulação de 100 cenários equiprováveis obtidos pela Simulação Sequencial Gaussiana, serão definidos o pior e o melhor cenário da distribuição espacial da contaminação, além da média das simulações para cada nó do grid, os quais serão analisados a partir de mapas de probabilidades gerados pela combinação das simulações calculadas.

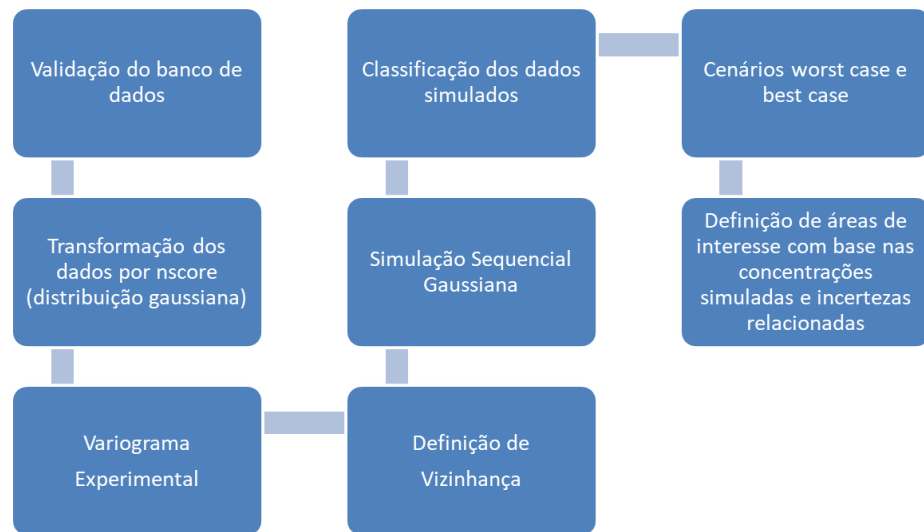
### 5.2. Métodos

De forma a atingir os objetivos propostos no presente estudo, serão aplicadas as etapas metodológicas conforme a sequência apresentada no fluxograma da Figura 6 e discutida no item 4.

Para o tratamento do banco de dados, execução da análise estatística e aplicação dos métodos geoestatísticos, será utilizado o software Isatis®, por meio da licença de uso disponibilizada pelo Instituto de Geociências da Universidade de São Paulo.

Adicionalmente, o software *Quantum Gis* será utilizado para elaboração de mapas após o processamento dos dados.

Figura 6 – Sequência de etapas metodológicas previstas para o trabalho.



Fonte: Autora

## 6. RESULTADOS E DISCUSSÃO

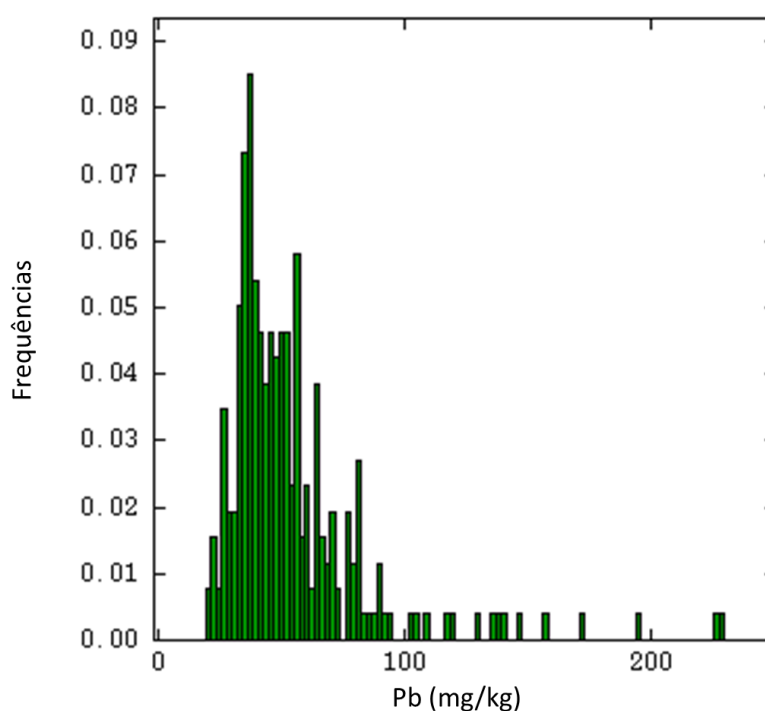
### 6.1. Análise Exploratória dos dados

A análise exploratória dos dados foi realizada no software Isatis® e contemplou as etapas de análise estatística dos dados brutos e transformados por *Normal Score*, e geoestatística dos dados transformados.

#### 6.1.1 Análise Estatística e Geoestatística dos dados

A partir do histograma da variável chumbo das 259 amostras do banco de dados, foi identificada uma distribuição assimétrica positiva, apresentando maiores frequências para baixas concentrações e poucos *outliers* de concentrações elevadas. O histograma e o mapa de pontos são apresentados nas Figuras 7 e 8, respectivamente, e a Tabela 1 apresenta as estatísticas descritivas da variável.

Figura 7 – Histograma da variável chumbo.



Fonte: Autora

Figura 8 – Mapa de pontos dos dados amostrais.

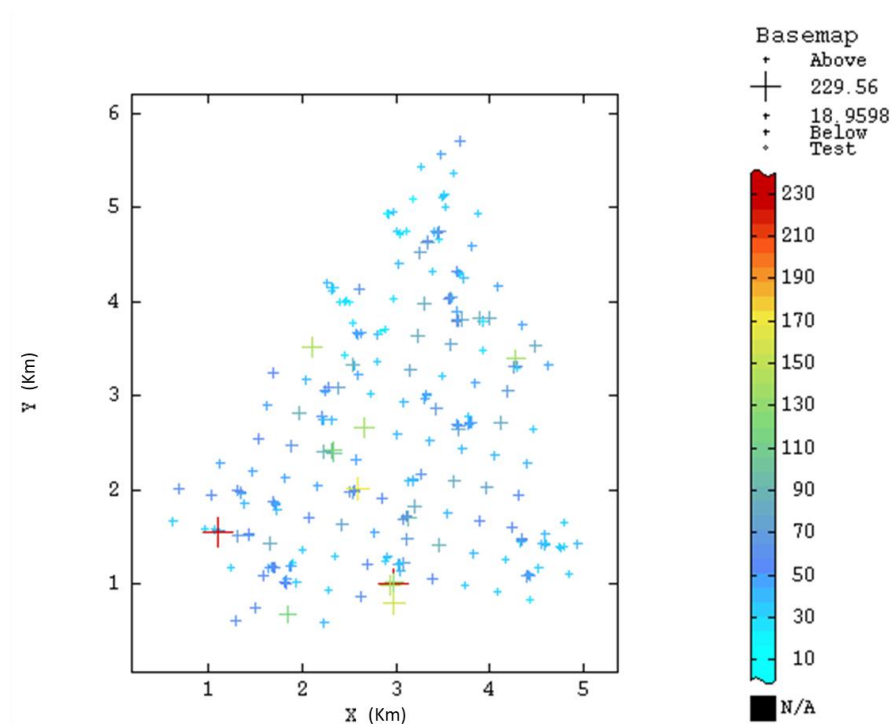


Tabela 1 – Estatística Descritiva dos dados de Chumbo (mg/kg).

| Nº de Amostras | Mínimo | 1º Quartil | Mediana | Média | 3º Quartil | Máximo | Desvio Padrão |
|----------------|--------|------------|---------|-------|------------|--------|---------------|
| 259            | 18,96  | 36,52      | 46,4    | 53,92 | 60         | 229,56 | 29,73         |

Fonte: autora

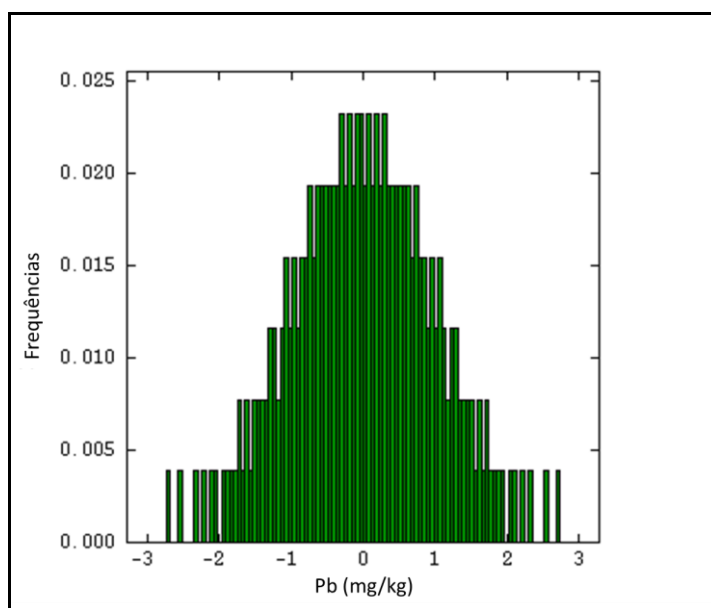
No presente trabalho, após a aplicação da SGS, os dados amostrais e simulados serão analisados com base no valor orientador de 50 mg/kg para chumbo, estabelecido pela Agência Nacional de Meio Ambiente da Suíça (FOEFL, 1989). Esta concentração foi utilizada como limite crítico em trabalhos como Atteia (1994) e Goovaerts et al. (1997) para caracterizar a contaminação no solo e probabilidade de risco na área de Jura. Adicionalmente, para fins de exercício e ilustração das possibilidades do método da SGS, foram utilizados também o valor de prevenção e o limite de intervenção agrícola da CETESB (2021) para chumbo no solo, correspondentes a 72 mg/kg e 150 mg/kg, respectivamente.

Verificou-se que 109 (42,08 %) dentre as 259 amostras analisadas apresentam concentrações acima do valor orientado de 50 mg/kg, sendo que a concentração média (53,92 mg/kg) também se encontra acima desse limite. Com relação ao valor de prevenção, de 72

mg/kg, 40 amostras (15%) apresentaram concentração igual ou superior a esse limite, já para o valor de intervenção para solos agrícolas, de 150 mg/kg, apenas 5 pontos amostrais (1,9%) apresentaram concentrações superiores.

A etapa de transformação dos dados pelo método de *Normal Score* foi realizada previamente à etapa de análise exploratória, conforme o procedimento descrito no item 4.3.2.1. Os dados foram transformados para uma distribuição simétrica, com média igual à zero e variância igual à 1 ( $N[0,1]$ ), permitindo assim a aplicação do método de Simulação Sequencial Gaussiana. As estatísticas descritivas e o histograma da distribuição dos dados transformados são apresentados na Tabela 2 e na Figura 9, respectivamente.

Figura 9 – Histograma da variável chumbo transformada por *nscore*.



Fonte: Autora

Tabela 2 - Estatísticas Descritivas dos dados de Chumbo transformados por *normal score*.

| Nº de Amostras | Mínimo | 1º Quartil | Mediana | Média | 3º Quartil | Máximo | Desvio Padrão |
|----------------|--------|------------|---------|-------|------------|--------|---------------|
| 259            | -2,74  | -0,68      | 0       | 0     | 0,68       | 2,74   | 0,99          |

Fonte: autora

Em seguida, o variograma experimental foi calculado com o objetivo de reconhecer as anisotropias existentes e verificar o comportamento espacial do fenômeno. Como a malha de

dados possui disposição espacial irregular e bidimensional, a análise exploratória ocorreu primeiramente em quatro direções: 0°, 45°, 90° e 135°.

Com base nos variogramas experimentais resultantes, foram escolhidas as direções 0° e 90° como eixos da elipse que representa a anisotropia, sendo que direção 0° corresponde ao maior eixo e a de 90° representa o menor eixo. A divergência de comportamento dos variogramas foi o critério de escolha dessas duas direções, sendo que a direção de 0° apresentou maior continuidade. Os parâmetros utilizados na análise exploratória são apresentados na Tabela 3.

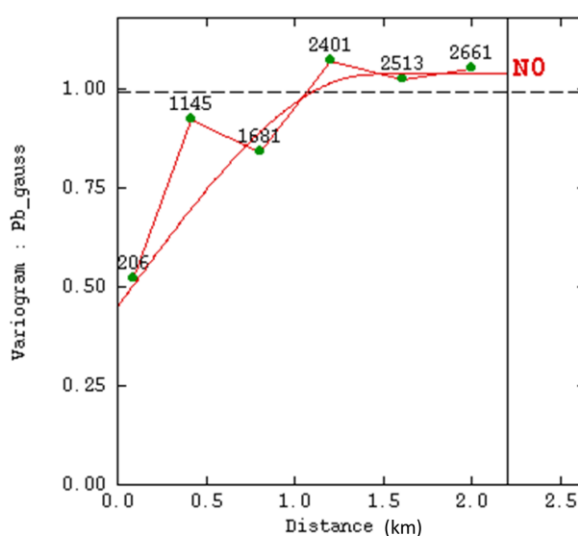
Tabela 3 – Parâmetros utilizados da análise exploratória.

| Direções | Passo | Nº de passos | Tolerância angular | Tolerância na distância |
|----------|-------|--------------|--------------------|-------------------------|
| 0° e 90° | 0,4   | 6            | 45°                | 0,2                     |

Fonte: Autora

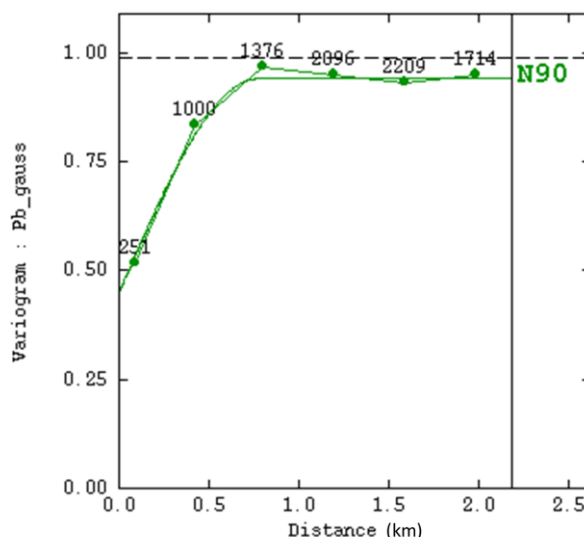
Posteriormente foi realizado o ajuste do modelo teórico do variograma, para o qual foi considerada anisotropia mista e modelo esférico. Os parâmetros de cálculo do variograma são apresentados na Tabela 4 e os variogramas experimentais e modelos de variograma ajustados para as direções de 0° e 90° podem ser observados nas Figura 10 e 11.

Figura 10 - Variograma experimental e modelo de variograma ajustado para a direção 0°.



Fonte: Autora

Figura 11 - Variograma experimental e modelo de variograma ajustado para a direção 90°.



Fonte: Autora

Tabela 4 – Estruturas dos modelos de variogramas experimentais

| Estruturas   | Direções | Efeito pepita | Modelo   | a máx | a mín   | Variância Espacial |
|--------------|----------|---------------|----------|-------|---------|--------------------|
| 1ª estrutura | 0°/0°    | 0,45          | Esférico | 1,4   | 0,77    | 0,49               |
| 2ª estrutura | 90°/0°   |               |          | 1,7   | 1,0E+30 | 0,1                |

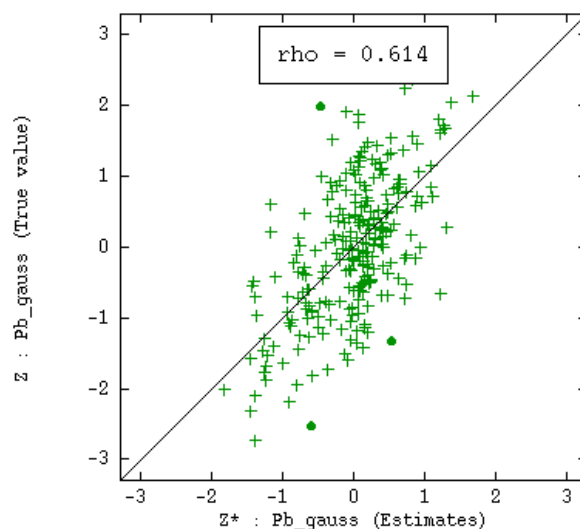
Fonte: Autora

## 6.2. Validação Cruzada

Após a modelagem do variograma, foi realizada a etapa de validação cruzada, durante a qual foram testados diversos parâmetros de vizinhança de forma a garantir que os dados simulados sejam representativos do conjunto de dados amostrais.

A etapa de validação cruzada foi realizada de acordo com o procedimento descrito no item 4.2. Os parâmetros de vizinhança que resultaram no maior coeficiente de correlação de Pearson, de 0,614, são apresentados na Tabela 5. A Figura 12 representa o diagrama de dispersão entre os dados amostrais e estimados para a variável chumbo, ambos em *normal score*.

Figura 12: Diagrama de dispersão entre os dados amostrais e estimados transformados por *normal score*.



Fonte: Autora

Tabela 5 – Parâmetros de busca aplicados na SGS.

| Parâmetros                         |     |
|------------------------------------|-----|
| Busca em X (km)                    | 1.3 |
| Busca em Y (km)                    | 0.5 |
| Número mínimo de amostras          | 4   |
| Número de setores horizontais      | 4   |
| Número ótimo de amostras por setor | 2   |

Fonte: Autora

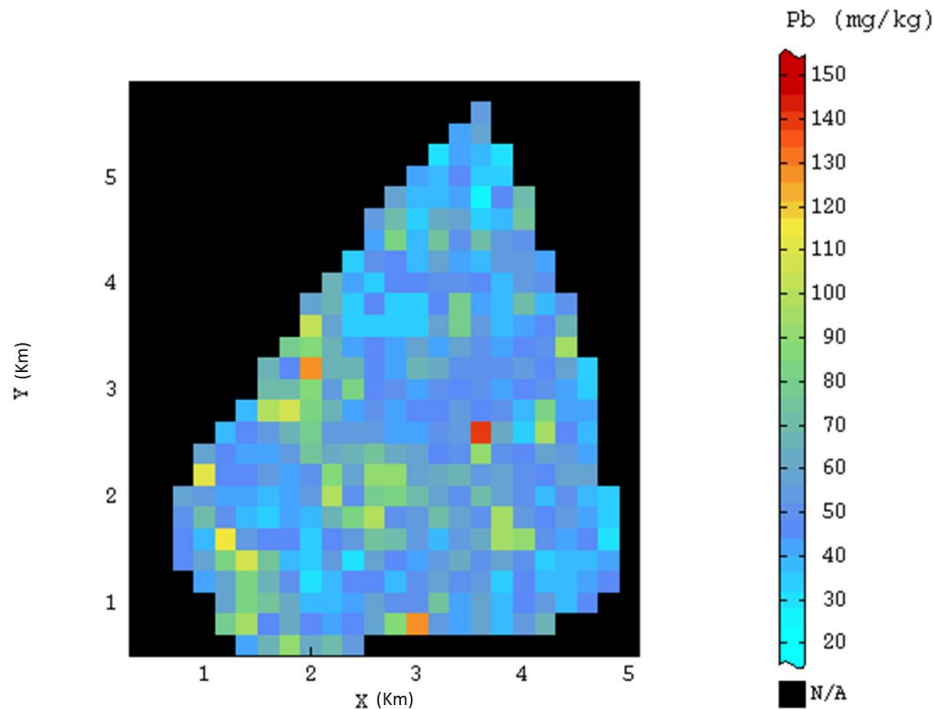
### 6.3.Simulação Sequencial Gaussiana

O método de SGS foi aplicado utilizando-se os parâmetros de vizinhança, a partir do qual foram geradas 100 realizações de igual probabilidade de ocorrência. Após a realização da simulação, os dados foram transformados novamente para o espaço original dos dados amostrais...

Com o objetivo de limitar a aplicação da SGS à área com presença de dados amostrais, foi gerada uma fronteira convexa, que consiste um polígono que abrange domínio amostral, o qual não é extrapolado durante a aplicação do método.

Na Figura 13, a título de ilustração, consta a realização nº 78, escolhida aleatoriamente dentre todos os resultados obtidos pela SGS

Figura 13: Realização n° 78 para a variável chumbo, resultante do método de SGS.



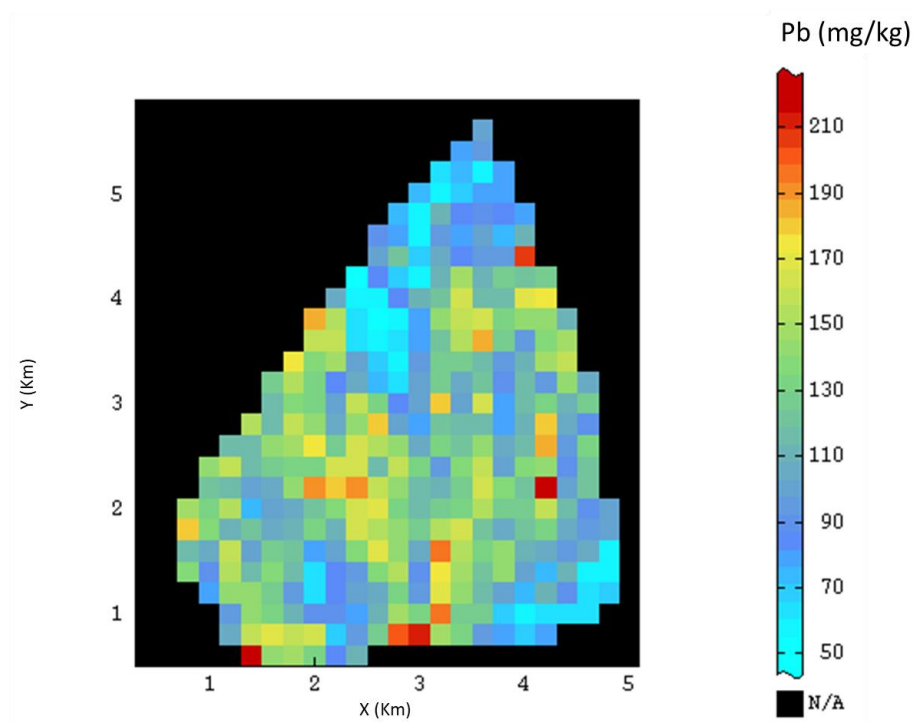
Fonte: Autora

### 6.3.1. Simulações de melhor e pior cenário

Para fins de exercício e ilustração, serão apresentados os mapas de pior e melhor cenário de distribuição das concentrações de chumbo no solo. Os mapas foram elaborados a partir do conjunto de valores máximos (pior cenário) e mínimos (melhor cenário) para cada bloco, considerando as 100 realizações geradas. Ressalta-se que tais cenários não correspondem a realizações específicas e sim a valores obtidos em diferentes realizações, sendo assim, a estatística descritiva e geoestatística dos dados não possuem as mesmas características dos dados amostrais e, portanto, não são realistas.

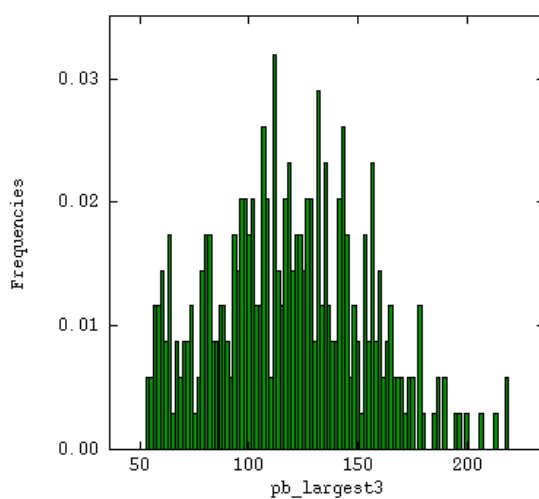
O conjunto de valores simulados que gerou o pior cenário, com as concentrações mais elevadas para chumbo para cada bloco, é apresentado na Figura 14. O histograma correspondente está representado na Figura 15 e as respectivas estatísticas descritivas encontram-se na Tabela 6.

Figura 14: Mapa de pior cenário para a variável chumbo



Fonte: Autora

Figura 15: Histograma dos dados simulados para o cenário de maiores concentrações de chumbo.



Fonte: Autora

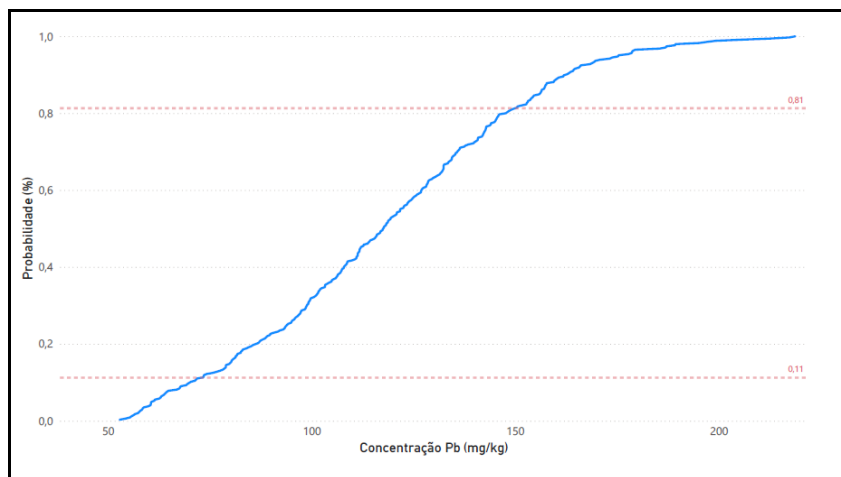
Tabela 6 – Estatísticas descritivas do cenário de maiores concentrações para chumbo.

| Nº de Amostras | Mínimo | 1º Quartil | Mediana | Média  | 3º Quartil | Máximo | Desvio Padrão |
|----------------|--------|------------|---------|--------|------------|--------|---------------|
| 345            | 52,89  | 94,27      | 117,74  | 118,29 | 142,55     | 218,77 | 34,53         |

Fonte: Autora

Adicionalmente, foi elaborado o gráfico de probabilidade de frequências acumuladas referente ao pior cenário, conforme apresentado na Figura 16. A partir da análise da curva de frequência, nota-se o valor de 72 mg/kg corresponde ao percentil 11% e o valor de 150 mg/kg corresponde ao percentil 81% da distribuição acumulada, como pode ser observado na Tabela 7.

Figura 16: Gráfico de distribuição de frequências acumuladas para os valores máximos simulados para cada nó.



Fonte: Autora

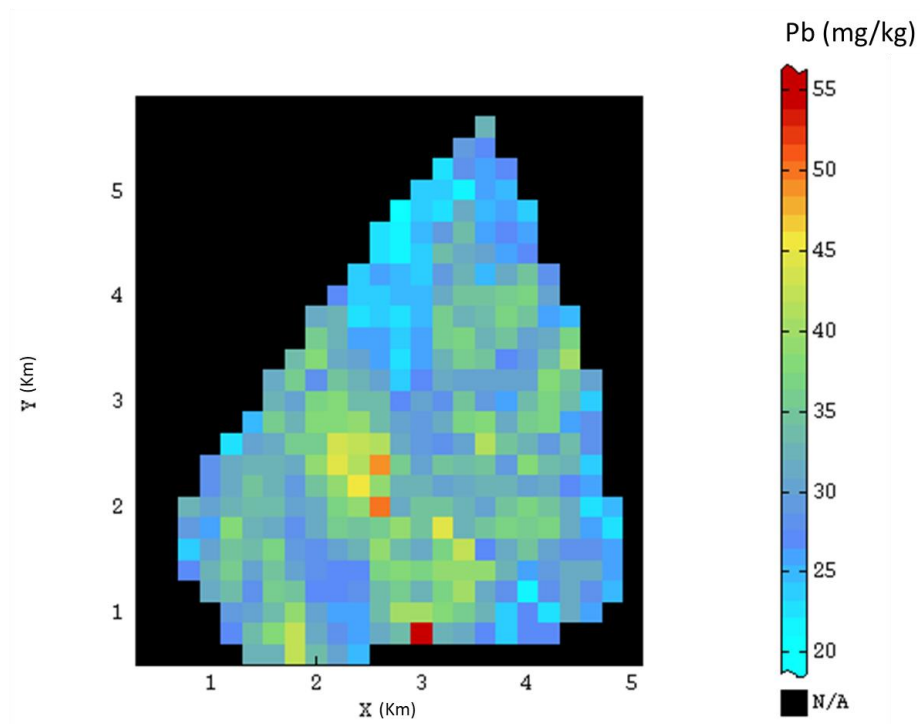
Tabela 7 - Percentis correspondentes aos valores de referência para o pior cenário simulado

| Valor Orientador<br>(mg/kg) | Frequencia<br>acumulada (%) | Área impactada (%) |
|-----------------------------|-----------------------------|--------------------|
| 50                          | -                           | 100                |
| 72                          | 11                          | 88,98              |
| 150                         | 81                          | 18,84              |

Fonte: Autora

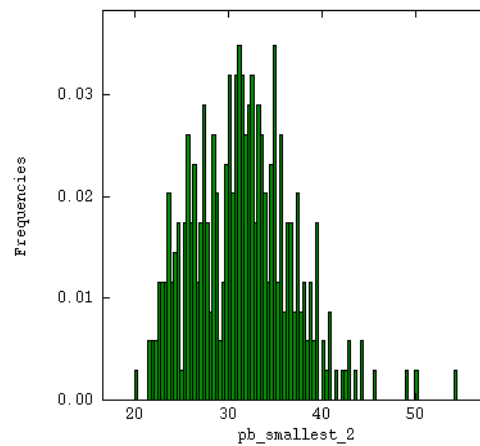
Já o melhor cenário, considerando os menores valores simulados para cada bloco, é apresentado na Figura 17. As estatísticas descritivas dos dados simulados encontram-se na Tabela 8 e o respectivo histograma pode ser observado na Figura 18. Observa-se que apenas 2 pontos simulados (0,01%) apresentaram valores maiores que 50 mg/kg. Nenhum valor simulado ultrapassou os valores orientadores de 72 mg/kg e 150 mg/kg neste cenário, como pode ser visto no gráfico de frequência acumulada da Figura 19 e Tabela 9.

Figura 17: Mapa de melhor cenário para a variável chumbo



Fonte: Autora

Figura 18 - Histograma dos dados simulados para o cenário de menores concentrações de chumbo.

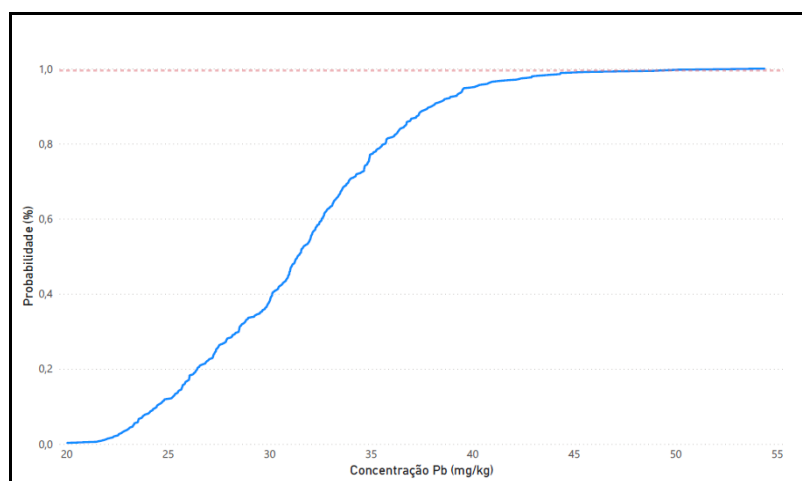


Fonte: Autora

Tabela 8 – Estatísticas descritivas do cenário de menores concentrações para chumbo.

| Nº de Amostras | Mínimo | 1º Quartil | Mediana | Média | 3º Quartil | Máximo | Desvio Padrão |
|----------------|--------|------------|---------|-------|------------|--------|---------------|
| 345            | 20,04  | 27,38      | 31,41   | 31,49 | 34,9       | 54,39  | 5,31          |

Figura 19: Gráfico de distribuição de frequências acumuladas para os valores mínimos simulados para cada nó.



Fonte: autora

Tabela 9 - Frequências acumuladas correspondentes aos valores de referência para o melhor cenário simulado.

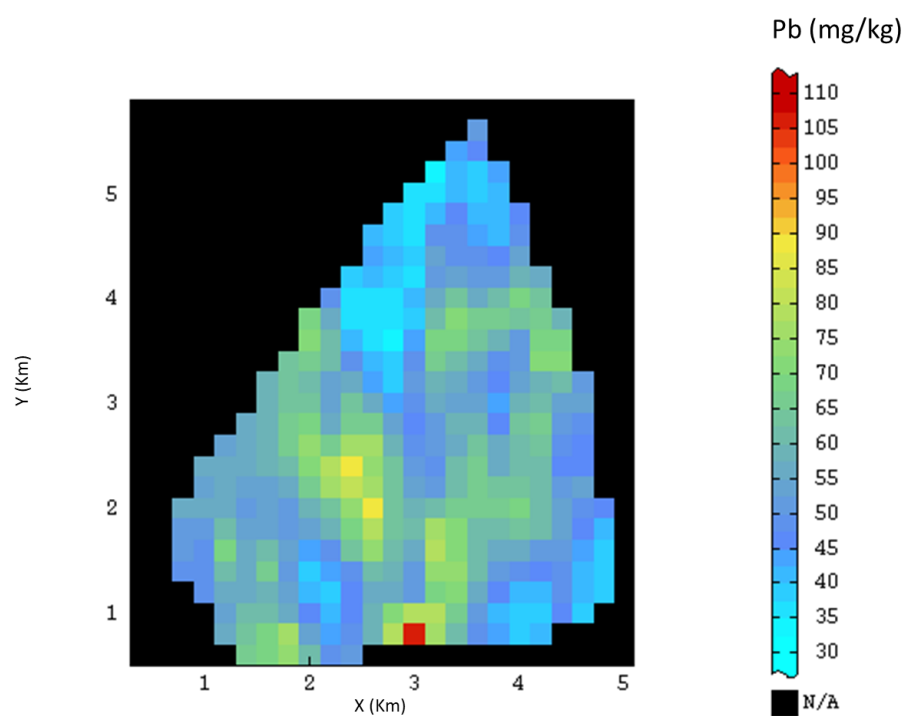
| Valor Orientador<br>(mg/kg) | Frequencia<br>acumulada (%) | Área impactada (%) |
|-----------------------------|-----------------------------|--------------------|
| 50                          | 1                           | 0,57               |
| 72                          | -                           | 0                  |
| 150                         | -                           | 0                  |

Fonte: Autora

### 6.3.2. Média do conjunto de realizações

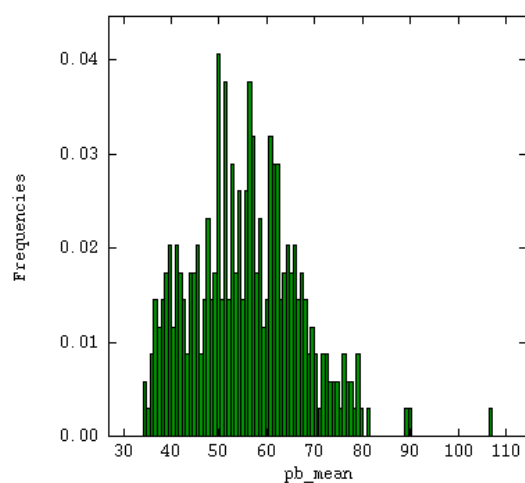
A partir da combinação das 100 simulações, foi gerado o mapa de concentrações médias, o qual é apresentado na Figura 20. O histograma e as estatísticas descritivas correspondentes às concentrações médias obtidas para cada um dos nós simulados do grid, considerando as 100 realizações, encontram-se na Figura 21 e Tabela 10, respectivamente.

Figura 20: Mapa de médias obtido a partir do conjunto de simulações.



Fonte: Autora

Figura 21: Histograma dos dados médios obtidos a partir do conjunto de realizações para a variável chumbo



Fonte: Autora

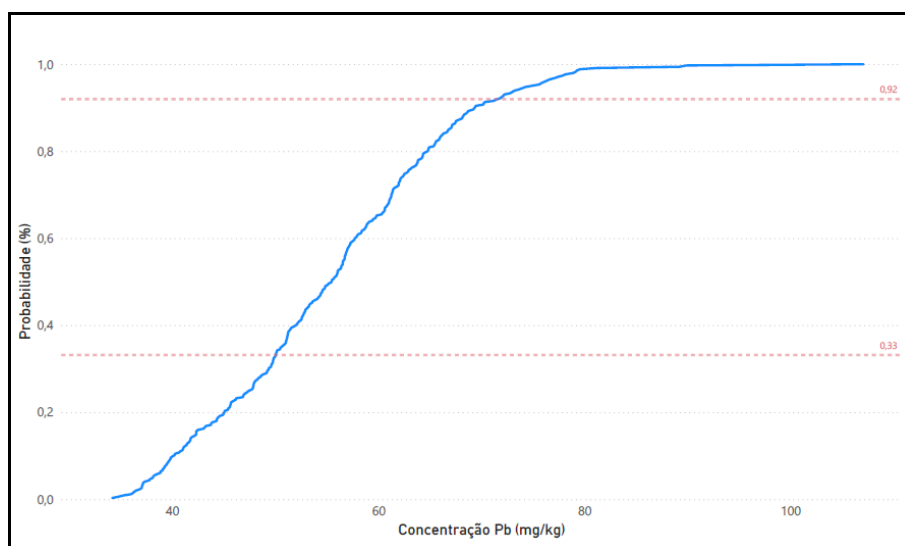
Tabela 10 - Estatísticas descritivas dos dados simulados médios.

| Nº de Amostras | Mínimo | 1º Quartil | Mediana | Média | 3º Quartil | Máximo | Desvio Padrão |
|----------------|--------|------------|---------|-------|------------|--------|---------------|
| 345            | 34,17  | 47,7       | 55,47   | 55,48 | 62,76      | 107,08 | 11,29         |

Fonte: Autora

Foram identificados 232 (67%) valores médios que ultrapassam o limite crítico de 50 mg/kg, equivalente ao percentil 33% da curva de distribuição de frequência acumulada. Para o valor de 72 mg/kg, correspondente a 92% da curva de distribuição de frequência acumulada, foram identificadas 26 (7,5%) médias superiores. Com relação ao valor de 150 mg/kg, todos os pontos apresentaram médias inferiores. O gráfico de distribuição de frequência acumulada pode ser observado na Figura 22 e os percentis correspondentes aos valores orientadores, encontram-se na Tabela 11.

Figura 22: Gráfico de distribuição de frequências acumuladas para as médias de valores simulados para cada nó.



Fonte: Autora

Tabela 11 – Percentis correspondentes aos valores de referência.

| Valor Orientador<br>(mg/kg) | Frequencia<br>acumulada (%) | Área impactada (%) |
|-----------------------------|-----------------------------|--------------------|
| 50                          | 33                          | 67,24              |
| 72                          | 92                          | 7,53               |
| 150                         | -                           | 0                  |

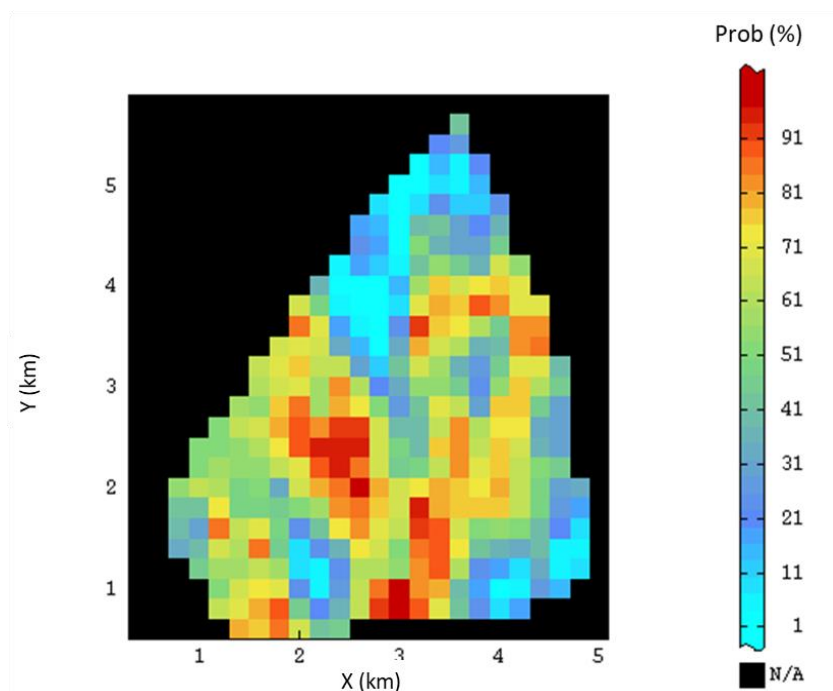
Fonte: Autora

### 6.3.3. Mapas de probabilidade de extrapolação dos valores orientadores

Com base no limite crítico de 50 mg/kg, correspondente ao valor orientador para concentrações de chumbo no solo (FOEFL, 1987), foi gerado o mapa de probabilidade ilustrado na Figura 23. A Tabela 12 apresenta as porcentagens da área de estudo com

probabilidades 50%, 75% e 95% de extrapolação do limite de 50 mg/kg, com base nos valores simulados.

Figura 23: Mapa de probabilidade para o valor de 50 mg/kg



Fonte: Autora

Tabela 12 – Porcentagens da área total com probabilidade de extrapolação do limite de 50 mg/kg.

| Probabilidade<br>> 50 mg/kg | Área impactada<br>(%) |
|-----------------------------|-----------------------|
| 95                          | 2,03                  |
| 75                          | 22,61                 |
| 50                          | 57,68                 |

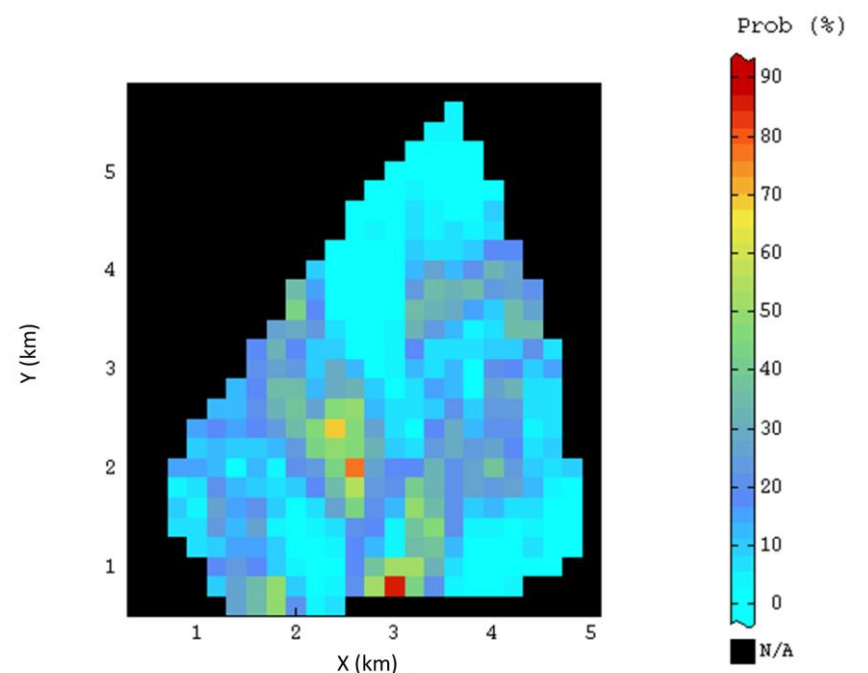
Fonte: Autora

Nota-se que 2% da área de estudo apresenta probabilidade acima ou igual à 95% de apresentar valores superiores a 50 mg/kg, 22,61% da área apresentou probabilidade acima ou igual à 75% e 57,68% apresentaram probabilidade igual ou acima de 50% de extrapolação do limite.

Para o limite de 72 mg/kg, correspondentes ao valor de prevenção da CETESB (2021) para concentrações de chumbo no solo, foi gerado o mapa de probabilidade de extrapolação apresentado na Figura 24. Conforme apresentado na Tabela 13, apenas 9 dos 345 nós simulados apresentam probabilidade de 50% ou superior de extrapolação do limite de 72 mg/kg, correspondente a 2,6% da área de estudo. Considerando 75% de probabilidade, a área

impactada passa a representar apenas 0,6% do total e nenhum dos pontos apresentou acima de 95% de probabilidade de extrapolação.

Figura 24: Mapa de probabilidade para o valor de 72 mg/kg



Fonte: Autora

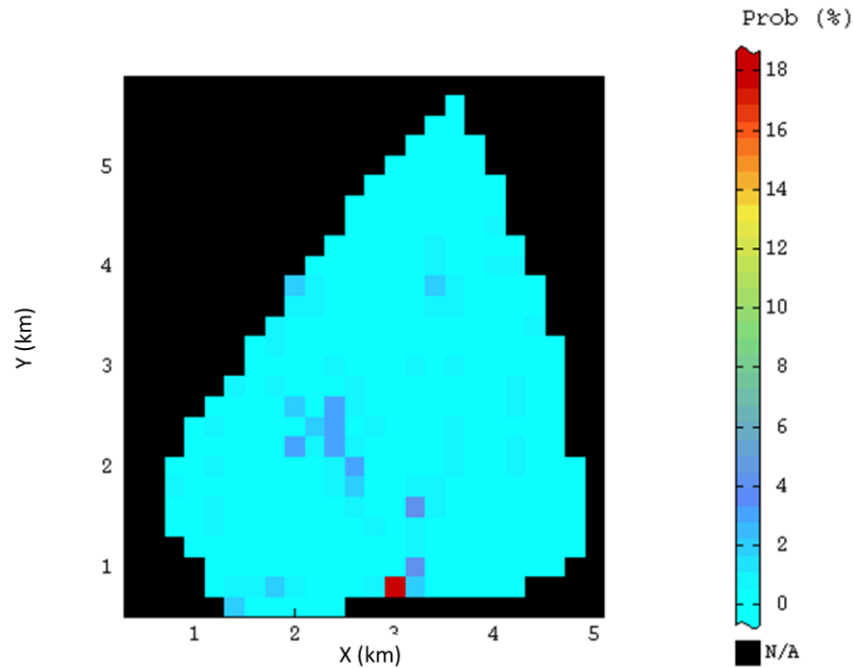
Tabela 13 – Porcentagens da área total com probabilidade de extrapolação do limite de 72 mg/kg.

| Probabilidade<br>> 72 mg/kg | Área impactada<br>(%) |
|-----------------------------|-----------------------|
| 95                          | 0                     |
| 75                          | 0,6                   |
| 50                          | 2,6                   |

Fonte: Autora

Para o limite de 150 mg/kg, correspondente valor de intervenção para solos agrícolas (CETESB, 2021), as probabilidades de extrapolação foram muito inferiores aos demais limites analisados, com probabilidade máxima de 18% observada em apenas um ponto. Os demais nós da malha apresentaram probabilidades entre 0 e 4 %. O mapa de probabilidade para 150 mg/kg é apresentado na Figura 25.

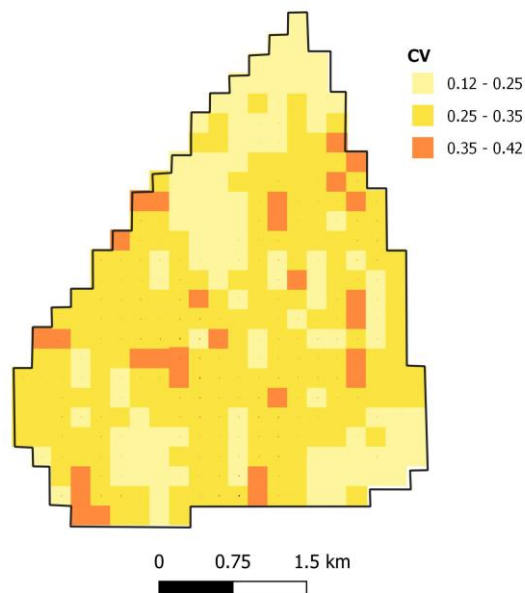
Figura 25: Mapa de probabilidade para o valor de 150 mg/kg



Fonte: Autora

A partir dos valores de média simulados e o desvio padrão para cada ponto, foi elaborado um mapa de coeficiente de variação, apresentado na Figura 26. Nota-se que todos os valores obtidos são baixos, sendo que o coeficiente máximo verificado foi de 0,42, o mínimo 0,12 e a média igual a 0,28. As magnitudes dos coeficientes de variação indicam que os valores simulados possuem baixa incerteza, visto que o coeficiente mais alto, de 0,42, é inferior a metade da média.

Figura 26 – Mapa de coeficientes de variação



Fonte: Autora

#### 6.3.4. Mapas de risco

Com base nos produtos resultantes da Simulação Sequencial Gaussiana, foram elaborados mapas de risco utilizando as médias das realizações, o desvio padrão de cada nó do grid, e o valor de corte de 50% probabilidades de extrapolação dos limites críticos, assumindo um cenário conservador e um não conservador.

O critério utilizado na identificação e classificação do risco para o cenário conservador foi a combinação das duas condições apresentadas abaixo.

##### **Condição 1:**

- **Classe 1:** (Média das simulações + desvio padrão) > 50 mg/kg
- **Classe 2:** (Média das simulações + desvio padrão) > 72 mg/kg
- **Classe 3:** (Média das simulações + desvio padrão) > 150 mg/kg

##### **Condição 2:**

- **Classe 1:** Probabilidade de extrapolação do limite de 50 mg/kg > 50%
- **Classe 2:** Probabilidade de extrapolação do limite de 72 mg/kg > 50%
- **Classe 3:** Probabilidade de extrapolação do limite de 150 mg/kg > 50%

Já o critério utilizado na classificação do risco para o cenário não conservador foi a combinação das seguintes condições.

##### **Condição 1:**

- **Classe 1:** (Média das simulações - desvio padrão) > 50 mg/kg
- **Classe 2:** (Média das simulações - desvio padrão) > 72 mg/kg
- **Classe 3:** (Média das simulações - desvio padrão) > 150 mg/kg

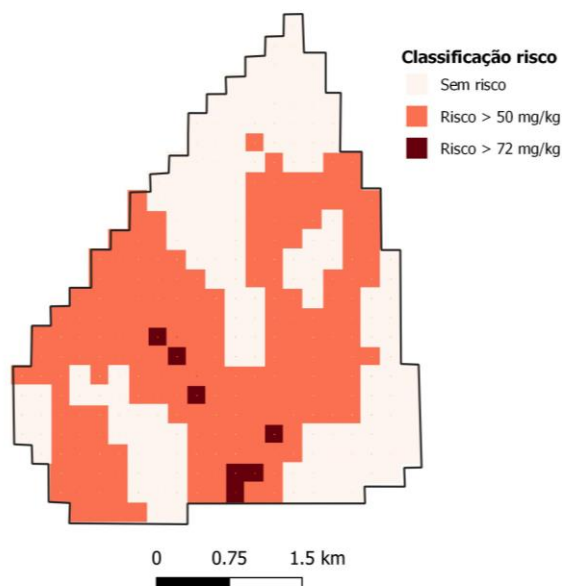
##### **Condição 2:**

- **Classe 1:** Probabilidade de extrapolação do limite de 50 mg/kg > 50%
- **Classe 2:** Probabilidade de extrapolação do limite de 72 mg/kg > 50%

- **Classe 3:** Probabilidade de extrapolação do limite de 150 mg/kg > 50%

O mapa resultante para o cenário conservador pode ser observado na Figura 27 e as porcentagens da área impactada para cada valor podem ser observadas na Tabela 14.

Figura 27 – Mapa de risco de extrapolação dos valores orientadores – Cenário conservador



Fonte: Autora

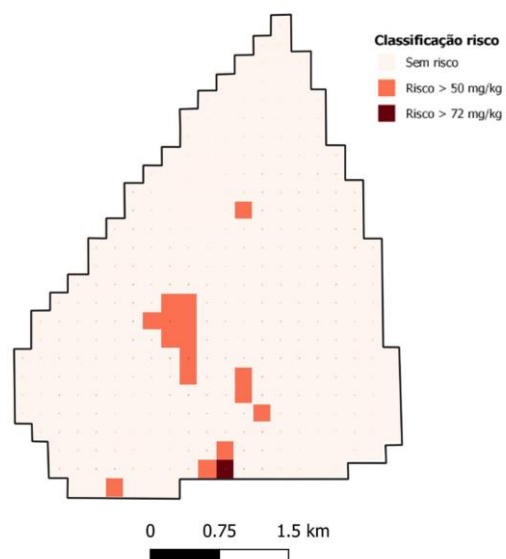
Tabela 14 – Porcentagens da área total com risco de extrapolação dos limites orientadores.

| Valor Orientador<br>(mg/kg) | Área impactada (%) |
|-----------------------------|--------------------|
| 50                          | 56                 |
| 72                          | 2                  |
| 150                         | 0                  |

O mapa referente ao cenário não conservador é apresentado na Figura 28 e as porcentagens da área impactada considerando os valores orientadores podem ser observadas na Tabela 15.

Verifica-se que em ambos os cenários, nenhum dos blocos foi atendeu às condições da classe 3 de risco, referente à extrapolação do valor de intervenção de 150 mg/kg (cenário residencial da CETESB, 2021). Dessa forma, esta classe também não foi exibida na legenda dos mapas.

Figura 28 – Mapa de risco de extrapolação dos valores orientadores – Cenário não conservador



Fonte: Autora

Tabela 15 – Porcentagens da área total com risco de extrapolação dos limites orientadores – Cenário não conservador.

| Valor Orientador<br>(mg/kg) | Área impactada (%) |
|-----------------------------|--------------------|
| 50                          | 4,9                |
| 72                          | 0,3                |
| 150                         | 0                  |

Fonte: Autora

## 7. CONCLUSÕES E CONSIDERAÇÕES FINAIS

O presente trabalho utilizou-se da técnica de Simulação Sequencial Gaussiana (SGS) a fim de caracterizar a distribuição da concentração de chumbo no solo, na região de Jura, na Suíça. A partir das 100 realizações geradas pela SGS foi possível realizar uma análise probabilística do risco de contaminação, utilizando-se como base três diferentes valores críticos.

Os produtos resultantes da presente análise exemplificam como o método da SGS pode atuar como uma ferramenta no gerenciamento de áreas contaminadas, principalmente na etapa de avaliação de risco, auxiliando a tomada de decisões à medida que os diferentes cenários simulados fornecem resultados representativos da variabilidade espacial dos fenômenos.

No âmbito do gerenciamento de áreas contaminadas, durante as etapas de investigação os resultados das simulações das concentrações das substâncias químicas de interesse podem ser comparados com valores orientadores, como os estabelecidos pela CETESB (2021), Resolução CONAMA 420 (2009), entre outros, de forma a auxiliar a locação de pontos de amostragem, ou ainda na delimitação das plumas de contaminação, incluindo no plano vertical.

No entanto, o método de SGS representa grande vantagem especialmente durante a tomada de decisões nas etapas de avaliação de risco, elaboração do plano de intervenção e monitoramento da eficácia de métodos de remediação. Neste contexto os valores simulados podem ser analisados com base nas concentrações máximas aceitáveis para cada cenário identificado para a área, ou com as metas de remediação definidas.

A utilização dos valores simulados, cenários de maiores e menores concentrações, além das médias obtidas, podem ser utilizados para auxiliar a tomada de decisões de acordo com o contexto da área, levando em conta os demais fatores utilizados na análise de risco à saúde humana (substância química, rotas de exposição, receptores, ambiente fechado e aberto, meio impactado, etc.) além dos custos para remediação.

## 8. REFERÊNCIAS BIBLIOGRÁFICAS

ATTEIA, O., DUBOIS, J.-P. AND WEBSTER, R. Geostatistical analysis of soil contamination in the Swiss Jura. *Environmental Pollution*, 86, 315±27, 1994.

ARMSTRONG, M. **Basic linear geostatistics**. Berlin: Springer, 1998.

BOENTE, C. et al. A coupled multivariate statistics, geostatistical and machine-learning approach to address soil pollution in a prototypical Hg-mining site in a natural reserve. *Chemosphere*, v. 218, p. 767-777, 2019.

BRASIL. Conselho Nacional do Meio Ambiente -CONAMA. Resolução no420, de 28 de dezembro de 2009. Dispõe sobre critérios e valores orientadores de qualidade do solo quanto à presença de substâncias químicas e estabelece diretrizes para o gerenciamento ambiental de áreas contaminadas por essas substâncias em decorrência de atividades antrópicas. *Diário Oficial da República Federativa do Brasil, Poder Executivo, Brasília, DF, 30 dez. 2009. Seção 1, 20p.*

BROOS, M.J.; AARTS, L.; VAN TOOREN C.F.; STEIN A. Quantification of the effects of spatially varying environmental contaminants into a cost model for soil remediation, *Journal of Environmental Management*, Volume 56, Issue 2, 1999.

CETESB. Companhia Ambiental do Estado de São Paulo. Decisão de Diretoria nº 125/2021/E: Valores Orientadores para Solo e Águas Subterrâneas no Estado de São Paulo, São Paulo, 2021

CHEN, Jian et al. A joint standard-exceeding risk assessment of multiple pollutants based on robust geostatistics with categorical land-use type data: A case study of soil nitrogen and phosphorus. *Environmental Pollution*, p. 118901, 2022.

DEUTSCH, C. V.; JOURNEL, A. G. **Geostatistical software library and user's guide**. New York: Oxford University Press, 1998.

FOEFL (Swiss Federal Office of Environment, Forest and Landscape) Commentary on the Ordinance relating to Pollutants in Soil (VSBo of June 9, 1986). FOEFL, Bern. 1987.

FRANSSEN, H., GÓMEZ-HERNÁNDEZ, J. 3D inverse modelling of groundwater flow at a fractured site using a stochastic continuum model with multiple statistical populations. *Stochastic Environmental Research and Risk Assessment* 16, 155–174. 2002. <https://doi.org/10.1007/s00477-002-0091-7>

GOMEZ-HERNANDEZ, J. J., AND X.-H. WEN. To be or not to be multi-Gaussian? A reflection on stochastic hydrology, *Adv. Water Resources.*, 21, 46– 61. 1998.

GÓMEZ-HERNÁNDEZ, J. J; SAHUQUILLO, A; CAPILLA, J.E. Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data—I. Theory. *Journal of Hydrology*, v. 203, n. 1-4, p. 162-174, 1997.

GOOVAERTS, P. **Geostatistics for Natural Resources Evaluation**. New York: Oxford University Press, 1997.

GOOVAERTS, P., WEBSTER, R. & DUBOIS, JP. Assessing the risk of soil contamination in the Swiss Jura using indicator geostatistics. *Environmental and Ecological Statistics* 4, 49– 64. 1997.

GOOVAERTS, P. Geostatistics in soil science: state-of-the-art and perspectives, *Geoderma*, Volume 89, Issues 1–2, 1999, Pages 1-45,

GOOVAERTS, P. Impact of the simulation algorithm, magnitude of ergodic fluctuations and number of realizations on the spaces of uncertainty of flow properties, *Stochastic Environ. Res. Risk Asses.* 13, 161– 182, 1999.

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

ISTOK, J. D., and RAUTMAN, C A. "Probabilistic assessment of ground-water contamination: 2. results of case study." *Ground Water*, vol. 34, no. 6, Nov.-Dec. 1996, pp. 1050. LING, Y; CHANG, T. **Geostatistical simulation and estimation of the spatial variability of soil zinc**. *Journal of Environmental Science & Health Part A*. A35. 327-347. 10.1080/10934520009376974; 2000.

LV, Jianshu; LIU, Yang. An integrated approach to identify quantitative sources and hazardous areas of heavy metals in soils. *Science of the Total Environment*, v. 646, p. 19-28, 2019.

MATHERON, G. **Les variables régionalisées et leur estimation**. Paris: Masson, 1965. 306 p.

METAHNI, Sabrine et al. Comparison of different interpolation methods and sequential Gaussian simulation to estimate volumes of soil contaminated by As, Cr, Cu, PCP and dioxins/furans. *Environmental pollution*, v. 252, p. 409-419, 2019. PYRCZ, M.J., DEUTSCH, C.V. Transforming Data to a Gaussian Distribution. In J. L. Deutsch (Ed.), *Geostatistics Lessons*. Retrieved from <http://geostatisticslessons.com/lessons/normalscore>; 2018.

QUEIROZ, J. C. B. Utilização da geoestatística na quantificação do risco de contaminação por metais pesados, na área portuária de Santana-Amapá. 2003. xii, 170 f. Tese (doutorado) - Universidade Estadual Paulista, Instituto de Geociências e Ciências Exatas, 2003. Available at: <<http://hdl.handle.net/11449/102938>>.

ROCHA, M. M.; MIRANDA, C. S.; Catalani, T.G.T.; Paranhos Filho, A.C. Geoestatística aplicada a dados ambientais. In: Antonio Conceição Paranhos Filho; Camila Leonardo Miotto; José Marcato Junior; Thais Gisele Torres Catalani. (Org.). **Geotecnologias em Aplicações Ambientais**. 2ed. Campo Grande: UFMS, 2016, v. 2, p. 213-222.

YAMAMOTO, J. K. 2001. Análise Geoestatística. In: YAMAMOTO, J. K., 2001, **Avaliação e Classificação de Reservas Minerais**. Editora da Universidade de São Paulo, São Paulo, p. 72-78.

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

WEBSTER, R., ATTEIA, O. AND DUBOIS, J.-P. Coregionalization of trace metals in the soil in the Swiss Jura. *European Journal of Soil Science*, 45, 205±18, 1994.