

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE GEOCIÊNCIAS

SIMULAÇÃO ESTOCÁSTICA DE CONTAMINAÇÃO DE CÁDMIO EM JURA,
SUÍÇA

Fernando Uchidomari de Sousa

Orientador: Prof. Dr. Marcelo Monteiro da Rocha

MONOGRAFIA DE TRABALHO DE FORMATURA
(TF-2012/16)

SÃO PAULO
2012

SIMULAÇÃO ESTOCÁSTICA DE CONTAMINAÇÃO DE CÁDMIO EM JURA, SUÍÇA

Fernando Uchidomari de Sousa

Orientador: Prof. Dr. Marcelo Monteiro da Rocha

MONOGRAFIA DE TRABALHO DE FORMATURA
(TF-12 / 16)

Aluno: _____

De acordo: _____

SÃO PAULO
2012

Agradecimentos

Primeiramente, agradeço a meus pais, Luiz Paulo Fernandes de Sousa e Regiana Mazuru Uchidomari de Sousa, e também aos meus avós, Feliciano Bastos de Sousa e Idalinda Fernandes de Sousa, por todas as condições que me deram em todos os momentos, nunca me deixando faltar nada para que eu pudesse chegar onde estou.

Agradeço também ao Professor Doutor Marcelo Monteiro da Rocha, que além de um tutor, mostrou-se por diversas vezes um grande amigo, relação que espero poder levar por toda a vida.

Pelo abrigo sem o qual eu dificilmente estaria me formando, serei eternamente grato às minhas tias Miyoko Yamazato e Íris Yuri Uchidomari.

Ao Doutor Lucas Warren e ao Professor Doutor Paulo César Boggiani, agradeço não apenas pelo aprendizado geológico, mas por me ajudar a refletir sobre a forma de se levar a vida profissional.

Pelos diversos ensinamentos, agradeço imensamente ao Sidharta Rezende, e Renan Kendy Mancio, não apenas pela amizade, mas por tantas vezes me ajudar e incentivar a ir além dos meus limites.

Agradeço também aos meus irmãos André e Rafael Uchidomari de Sousa, por terem me ajudado em incontáveis aspectos, muitas vezes sem nem perceber.

Pela contribuição inestimável a este trabalho, agradeço a Stephano Silva, Eduardo Takafuji e, principalmente, a um dos meus melhores amigos, Felipe Casas.

Aproveito também para agradecer à Dayse Mantovani, ao Danilo Saunite, e também Bruna Pioli, Flávia Elise Decloedt, Daniela Borges, Malena D'Elia Otero, Letícia Freitas Guimarães, Priscila Cury e Cristiano Padalino Galeazzi, que tanto participaram da minha vida durante estes anos.

Agradeço também, como não podia deixar de ser, aos outros Lazarentos do Grupo VIII, Matheus Andrade, Raul Meloni e Gustavo Usero, com quem eu nunca achei que fosse rir tanto, mesmo com fome, sujo e de mau-humor.

E por último, mas não menos importante, agradeço por tornarem estes os melhores anos de minha vida até agora (apesar de alguns danos causados ao fígado, à face, à dignidade e ocasionalmente a calçadas) aos meus amigos André Polistchuck, Marcel Müller, Dhaniel Antunes de Carvalho, Elton Alves Marcelino, Gustavo Haruki Saka, Lucas Blanes de Oliveira, Leonardo Aguiar, Lucas Villela Cassini.

Existem muitas outras pessoas a quem eu gostaria de agradecer simplesmente por participarem da minha vida, mas o espaço que tenho para isso é muito pouco, infelizmente.

Índice

Resumo	1
Abstract	2
1-Introdução	3
2-Objetivos	4
3-Revisão Bibliográfica.....	5
3.1 – Base de dados.....	5
3.2 – Análise Estatística	7
3.3 – Análise Geoestatística	8
3.4 – O Modelo Multigaussiano	11
3.5 – Transformação Gaussiana	12
3.6 – Verificação da Normalidade Bivariada	14
3.7 – Simulação Estocástica.....	15
3.7.1 – A Simulação Sequencial Gaussiana	17
4-Tratamento dos dados	19
4.1 – Análise estatística dos dados	19
4.2 – Análise geoestatística dos dados não transformados.....	22
4.3 – Transformação dos dados	25
4.4 – Análise Geoestatística dos dados transformados.....	28
4.5 – Definição do <i>Grid</i> de Simulação	29
4.6 – A aplicação da simulação sequencial Gaussiana	30
4.6.1 – Pós-simulação.....	31
5-Conclusões	34
6-Referências Bibliográficas.....	36

Resumo

A geoestatística vem sendo cada vez mais utilizada no mapeamento de contaminantes e análises de risco de sítios impactados. Ao se observar as estimativas realizadas através da krigagem, porém, é possível perceber o efeito de suavização e autores como Olea (1999) apontam a simulação estocástica para se contrapor ao efeito de suavização, e neste caso foi aplicada a simulação sequencial Gaussiana (ssG), em que a variável é simulada sequencialmente em cada ponto desejado de acordo com a sua distribuição condicional local de probabilidade cumulativa (dclp). Este trabalho de formatura teve como objetivo treinar e familiarizar o aluno na aplicação de conceitos e ferramentas de geoestatística, voltando-os principalmente ao mapeamento de plumas de contaminação e análise de risco. Para tal, escolheu-se a base de dados fornecida em Goovaerts (1997), composta por ampla gama de dados de cunho ambiental, obtidos em Jura, Suíça.

Jura é uma região Suíça de atividade industrial relativamente reduzida, e em diversas localidades o solo apresentava elevado teor de Cd entre outros metais pesados. Apesar de sua grande empregabilidade tecnológica, o Cádmio é um elemento perigosamente tóxico para os seres humanos, embora seus danos sejam perceptíveis principalmente a longo prazo. Tratando-se de um caso de possível contaminação, os valores muito elevados foram analisados em detalhe e notou-se que boa parte dos valores anômalos está relativamente concentrada em alguns pontos do mapa evidenciando áreas com possíveis liberação de contaminantes.

Através dos variogramas direcionais, foi possível observar, para os dados originais e normalizados (obtidos pela função de anamorfose Gaussiana), que a elipse de anisotropia ocorre com o maior eixo a 90° e o menor a 0° .

Foram feitas 500 realizações da ssG, e de modo geral as áreas a noroeste e levemente a sul-sudeste do centro do *grid* (estendendo-se em algumas realizações para leste e oeste) apresentaram maiores incidências de concentrações elevadas de Cd.

Abstract

Geostatistics has been increasingly used in the mapping of contaminants and risk analysis of impacted sites. By observing the estimates made by kriging, however, one can see the smoothing effect and authors such as Olea (1999) point to the stochastic simulation to counteract the effect of smoothing, and in this case study the sequential Gaussian simulation (SGS) wherein the variable is simulated at each point sequentially according to local conditional probability distribution (lcpd). This study aimed to train and familiarize the student in the application of concepts and tools of geostatistics, especially to the mapping contamination plumes and risk analysis. In order to do so, the database provided in Goovaerts (1997) was chosen, comprising environmental nature data, obtained in Jura, Switzerland.

The Swiss Jura is a region of relatively low industrial activity, and in several places the soil presented high levels of Cd and other heavy metals. Despite its great technological employability, Cadmium is an dangerously toxic element to humans, though its damage is noticeable especially in the long term.

Since this is a case of possible contamination, very high values were analyzed in detail and it was noticed that most of the anomalous values are relatively concentrated in some points of the map showing areas with possible pouring of contaminants.

Through the directional variograms, it was possible to observe, for the original and normalized data (obtained by Gaussian anamorphosis function), the anisotropy ellipse with major axis occurs at 90° and at 0° , the lower.

500 realizations were made through the logarithm of SGS, and generally the areas slightly northwest and south-southeast of the grid center (extending in some executions east and west) presented higher incidences of elevated concentrations of Cd.

1- Introdução

A preocupação e o interesse com contaminações e recuperação de solos, recursos hídricos e o próprio ar cresce acentuadamente, um reflexo da conscientização ambiental adotada em grande parte do mundo. Um dos tipos principais de contaminantes do solo é o grupo dos metais pesados. Seja por infiltração de água através de aterros de lixo, rejeito em minerações, ou até mesmo da produção de combustíveis, elementos como Pb, Ni, Cd, Cu, Hg entre diversos outros, tem sido importantes alvos nos estudos de impacto ambientais de origem atrópica.

Quimicamente, os metais pesados são qualquer elemento que apresente propriedades metálicas, porém em geral é um termo utilizado para definir algum tipo de contaminante, embora nem todos os metais pesados sejam tóxicos (Hogan, 2010).

O cádmio, descoberto por Friedrich Stromeyer em 1817 (WEEKS & LEICESTER, 1968), é usado na fabricação de baterias (em associação com níquel), tintas, plásticos, diversos tipos de liga e em revestimentos anticorrosivos, sendo que esta última finalidade foi a mais comum, chegando a 65%, hoje sendo superada pelo uso em baterias portáteis (KRAFT, 1990).

Apesar de sua grande empregabilidade tecnológica, o Cádmio é um elemento perigosamente tóxico para os seres humanos, embora seus danos sejam perceptíveis principalmente a longo prazo. Um exemplo mundialmente conhecido se iniciou por volta de 1910, trata-se do caso da doença Itai-itai, que acometeu diversos agricultores e pescadores na região de Funchumachi, no Japão, devido a uma contaminação em larga escala causada pela extração e usina de processamento de Pb e Zn de companhias mineradoras, cujos processos liberaram grande quantidade de Cd no solo. Este Cd atingiu o rio Jintsu, envenenando peixes, e também plantações para as quais a água do rio era utilizada como irrigação. O cádmio foi ingerido pelos cidadãos de povoados próximos, de modo que uma parcela considerável destes começou a apresentar um padrão patológico. Um dos traços mais característicos desta doença é o enfraquecimento da estrutura óssea, a ponto de fraturas ocorrerem com muita facilidade, o que piora ainda mais a já forte dor causada pela doença, que inclusive é a origem de seu nome (Itai-itai é basicamente uma exclamação de dor no Japão).

Por mais óbvio que pareça atualmente, a doença, bem como suas causas, só foram reconhecidas em 1968, quando diversas vítimas abriram um processo contra a Mitsui Mining & Melting, a então responsável pelas atividades mineiras no local. Porém, apesar de declarada culpada e do impacto ter sido tratado pelas autoridades locais, o quadro de liberação de Cádmio no meio ambiente em escala mundial teve poucas melhoras (Friberg et al., 1971)

Um exemplo brasileiro a ser citado é o caso de Santo Amaro, município baiano, localizado nos arredores de extrações metalúrgicas de chumbo das décadas de 50 a 90. Juntamente com diversos outros contaminantes, o cádmio foi liberado na atmosfera por chaminés, contaminando solo e a água local. Tavares (1990) apontou teores perigosos de cádmio no sangue de crianças da região.

No solo o cádmio é relativamente estável (Eggenberger & Waber, 1988, NUS Corporation, 1987), porém pode ser transferido ao organismo do ser humano através de certos alimentos, cultivados em regiões de concentrações anômalas deste elemento.

De acordo com a OMS (2000), a exposição ao cádmio a médio e longo prazo tende em geral a afetar primeiramente os sistemas excretores, devido a sua acumulação nos rins.

2- Objetivos

Um dos grandes motivos de preocupação é que, de acordo com Borrough (1993), a organização espacial de elementos tóxicos no solo não é uniforme, uma vez que a distribuição não ocorre de forma homogênea através de diferentes camadas de solo e rocha. Deste modo torna-se evidente a importância de se mapear o comportamento dos elementos de interesse, para permitir a melhor ação de remediação.

A geoestatística vem sendo cada vez mais utilizada no mapeamento de contaminantes e análises de risco de sítios impactados. O desenvolvimento deste trabalho de formatura tem como objetivo treinar e familiarizar o aluno na aplicação de conceitos e ferramentas de geoestatística, voltando-os principalmente ao mapeamento de plumas de contaminação e análise de risco. Para tal, escolheu-se a base de dados fornecida em Goovaerts (1997) que é composta por ampla gama de dados de cunho ambiental, obtidos em Jura, Suíça (Figura 1).

Jura é uma região Suíça de atividade industrial relativamente reduzida, e foi percebido que em diversas localidades o solo apresentava elevado teor de Cd entre outros metais pesados. Este cantão, assim como todos os outros que compõe o país, foi incumbido de mapear o mais detalhadamente possível estes contaminantes. Neste âmbito diversos trabalhos foram desenvolvidos nesse âmbito, dos quais alguns focaram o emprego da geoestatística, como Atteia et al. (1994), que dimensionaram volumes deste e de outros metais utilizando principalmente krigagem ordinária.



Figura 1: Os três distritos do Cantão de Jura, na Suíça.

Este trabalho de formatura é um estudo deste caso e pretende-se realizar a análise geoestatística e posterior simulação estocástica, base para a análise de risco do sítio contaminado, levando em conta valores de referência utilizados em estudos geológicos ambientais.

3- Revisão Bibliográfica

3.1 – Base de dados

Os dados foram extraídos de Goovaerts (1997) e foram levantados e estudados inicialmente por Atteia et al. (1994) e Webster et al. (1994). São compostos por análises de metais pesados (como cádmio, chumbo, níquel e zinco), informações geológicas e de uso do solo (Goovaerts, 1997) coletadas em 359 pontos, como visto na Figura 2. As amostras foram

divididas em dois conjuntos, o de validação e o de predição, sendo que o primeiro consiste de amostras voltadas à confirmação das estimativas feitas com o segundo.

O conjunto de predição é composto por 259 amostras (Figura 2) distribuídas em malha regular com 107 pontos amostrais à distância de 250 m. O restante da amostragem teve por base 38 pontos escolhidos aleatoriamente. Partindo-se de cada um desses 38 pontos, tomou-se 100m de distância, em uma direção aleatória, para nova amostragem. A 40m de distância destas, também em direções aleatórias, foram coletadas mais 38 amostras, e assim subsequentemente às distâncias de 16 e 6m, compondo assim o conjunto de predição.

A amostragem foi feita desta forma, pois tratava-se de um caso de contaminação de fontes até então desconhecidas, de modo que qualquer possibilidade deveria ser averiguada, uma vez que não se sabia se as concentrações elevadas de cádmio e outros metais pesados poderiam estar associadas a fatores naturais ou antrópicos.

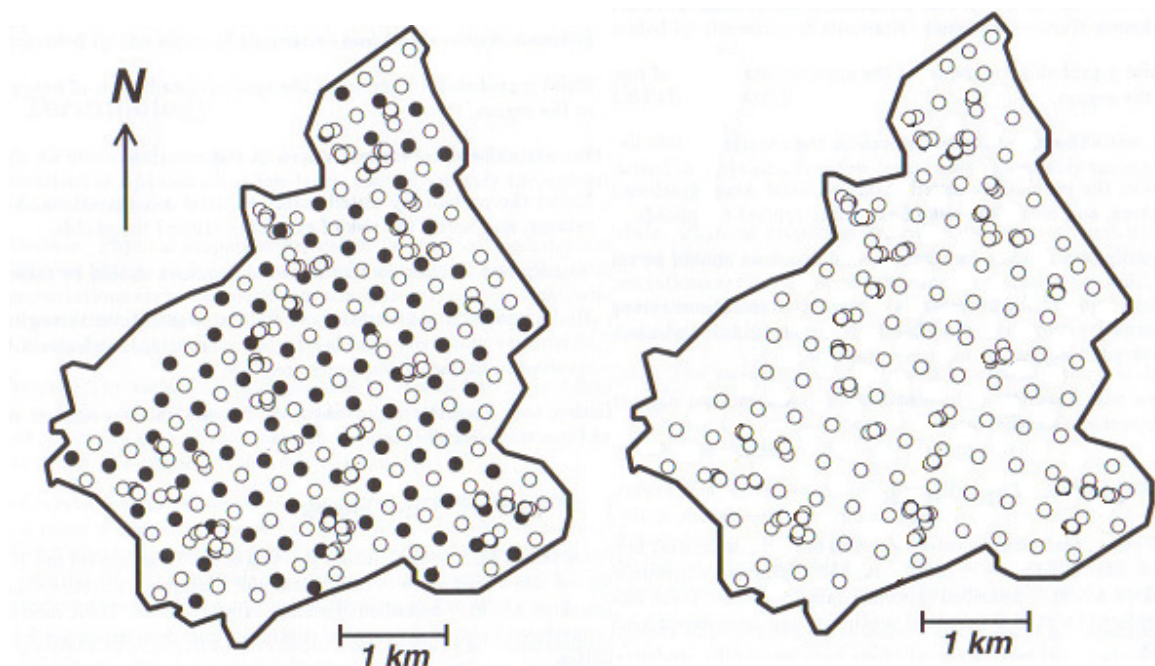


Figura 2: Mapa base das amostras coletadas em Jura, com todos os pontos amostrados à esquerda e o conjunto de predição à direita.

3.2 – Análise Estatística

A análise estatística é realizada anteriormente à análise geoestatística, visando a melhor compreensão dos dados, e permite a obtenção de medidas de tendência central, bem como a média e medidas de dispersão como a variância (Yamamoto, 2001).

De acordo com Yamamoto (2001) a distribuição de probabilidades pode ser apresentada de forma simples ou acumulada, podendo a primeira ser observada através do histograma, em que se gera um diagrama lançando intervalos de valores em abscissa e as frequências em ordenada (Figura 3), e na segunda, as frequências dos dados, agora são acumuladas.

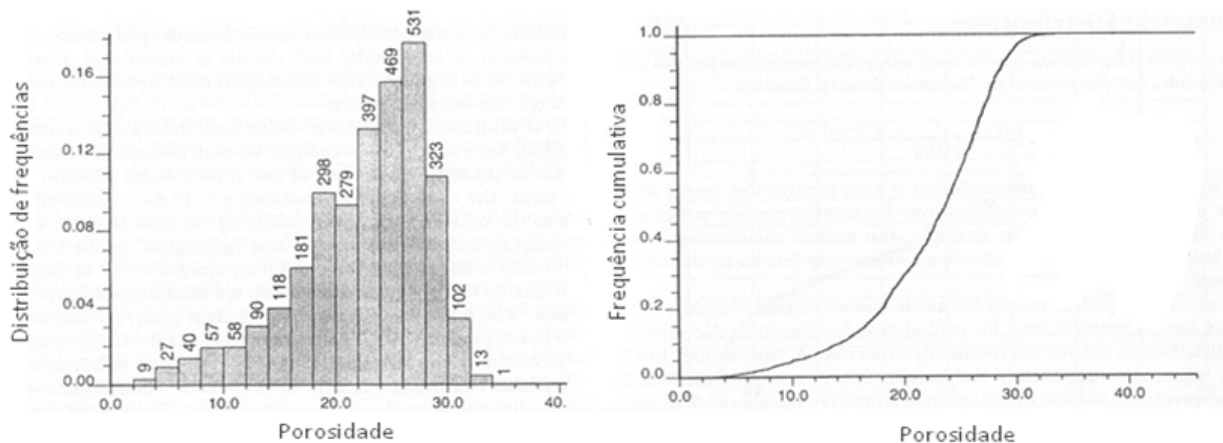


Figura 3: Exemplos de histograma e curva cumulativa de uma amostragem para uma medição de porosidade. Extraída de Deutsch, 2002.

As distribuições podem ser simétricas ou assimétricas (Figura 4), sendo o primeiro caso também denominado de distribuição normal ou gaussiana. Este trata-se do tipo mais utilizado, uma vez que é matematicamente conveniente, já que suas propriedades são bem conhecidas.

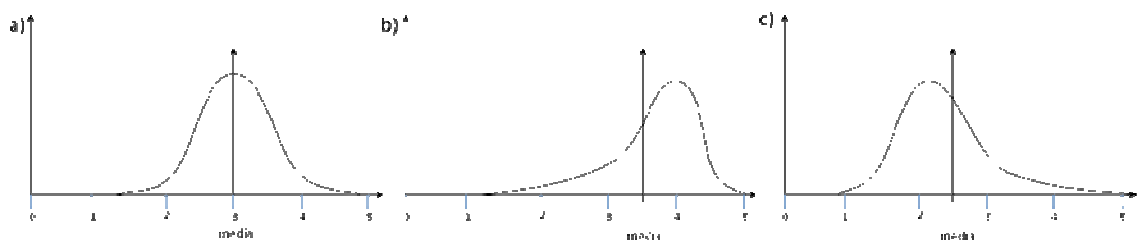


Figura 4: Curvas de distribuição de frequências simples, seguindo: a) uma distribuição normal (simétrica, gaussiana); b) uma assimetria negativa (truncada à direita) e; c) uma assimetria positiva (truncada à esquerda); (Yamamoto, 2001)

No caso das distribuições assimétricas, é possível se tratar tanto de uma assimetria positiva, em que são observados diversos valores baixos, e alguns poucos valores altos

(Yamamoto, 2001), quanto de assimetria negativa, que é o exato oposto, ou seja, um grande volume de valores elevados e poucas amostras com teores reduzidos

3.3 – Análise Geoestatística

Segundo Yamamoto (2001), nesta etapa obtém-se a ferramenta fundamental para a realização de estimativas, e simulações estocásticas, o variograma experimental. Este auxilia na compreensão da variabilidade espacial natural dos dados e a estimar/simular valores para pontos não amostrados.

O conceito de função aleatória é explicado em Deutsch & Journel (1992) como um conjunto de variáveis aleatórias, em geral relacionadas ao mesmo atributo, sobre o campo de interesse, denominado como Z . A variável aleatória é caracterizada pela sua função de distribuição cumulativa e, da mesma forma, uma função aleatória é caracterizada por todas as funções de distribuição cumulativa K -variadas, para qualquer número de amostragem K , em quaisquer K pontos de levantamento.

Assim como a função de distribuição cumulativa é usada para modelar a incerteza sobre algum valor dessa variável aleatória, uma função de distribuição cumulativa multivariada ($K < 1$) é usada para se modelar a incerteza conjunta dos K valores $z(u), \dots, z(u_k)$, onde z é o resultado de uma variável aleatória Z e u , de sua localização. Uma forma particularmente interessante de abordar tal característica é fazendo uso da função de distribuição cumulativa para $K=2$, ou seja, bivariada, de modo a se trabalhar com duas variáveis aleatórias. Deste modo é possível se obter como se dá a função distribuição de frequências entre duas variáveis aleatórias $Z(u)$ e $Z(u')$, nos pontos u e u' , chegando-se à covariância entre estas (Deutsch & Journel, 1992).

Ainda de acordo com Deutsch & Journel (1992), uma alternativa à covariância trata-se do covariograma, normalmente mais utilizado, definido como a variância do incremento $[Z(u) - Z(u+h)]$. Ou seja, para uma função aleatória estacionária:

$$2\gamma(\mathbf{h}) = \text{Var} \{Z(\mathbf{u} + \mathbf{h}) - Z(\mathbf{u})\}$$

$$\gamma(\mathbf{h}) = C(0) - C(\mathbf{h}), \forall \mathbf{u}$$

Sendo $C(h)$ a covariância estacionária e $C(0) = \text{Var} \{Z(u)\}$, a variância estacionária.

Com o variograma é possível observar se existe anisotropia, analisar a variabilidade espacial a curtas distâncias (através do comportamento próximo à origem), bem como identificar distâncias limítrofes para se admitir a influência entre amostras.

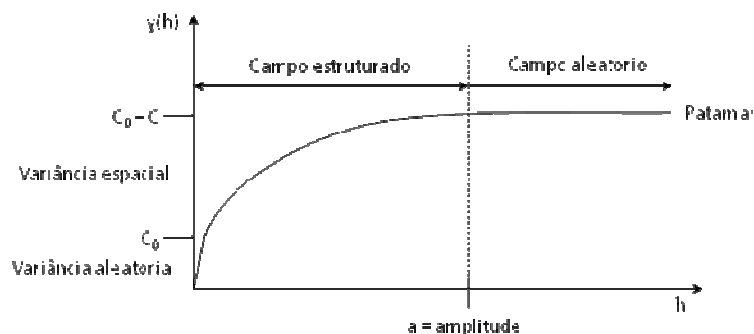


Figura 5: Variograma típico com suas propriedades (Yamamoto, 2001).

Como visto em Yamamoto (2001), podemos identificar no variograma várias características (Figura 5), dentre as quais, uma das mais importantes, é a amplitude, que é a distância a partir da qual as amostras já não tem nenhuma influência umas sobre as outras. O campo estruturado é o intervalo em que a amplitude ainda não foi atingida, ou seja, as amostras tem influência entre si, é o inverso do campo aleatório. O patamar indica o valor de variância no qual o variograma se estabiliza. O efeito pepita, também denominado como variância aleatória, é o valor da função variograma próximo à origem. A variância espacial, portanto, é calculada através da diferença entre patamar e o efeito pepita.

No variograma é possível observar o comportamento próximo à origem. Este é diretamente relacionado à continuidade da variável regionalizada a curtas distâncias. De acordo com Journel & Huijbregts (1978), o comportamento pode ser parabólico, linear, ou apresentar efeito pepita.

Como pode ser visto na Figura 6, o comportamento parabólico indica uma continuidade de alto grau entre as amostras a pequenas distâncias, enquanto que o linear apresenta, em distâncias muito pequenas, continuidade menor. O efeito pepita, por sua vez indica a descontinuidade, ou seja, indicariam que mesmo que a distância entre as amostras tenda a zero, a variância não será nula (Chillés e Delfiner, 1999). O efeito pepita pode se apresentar ainda como uma simples descontinuidade próximo à origem ou como efeito pepita puro, sendo este um caso particular onde as amostras não apresentam relação de dependência espacial para qualquer distância de observação.

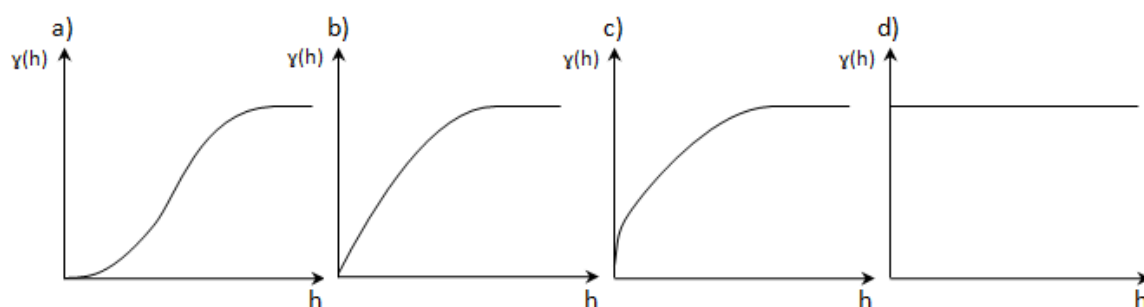


Figura 6: Comportamento próximo à origem de variogramas visando exemplificar: a) alta continuidade, parabólico; b) média continuidade, linear; c) descontinuidade como efeito pepita e; d) descontinuidade como efeito pepita puro (Yamamoto, 2001).

Há anisotropia quando variogramas em uma mesma área de estudo, tomados a diferentes direções, apresentam comportamentos diferentes. Segundo Yamamoto (2001), a anisotropia pode ser geométrica, zonal ou mista. A primeira é caracterizada por patamares idênticos, porém com amplitudes diferentes para cada direção. A anisotropia zonal ocorre quando os variogramas apresentam mesma amplitude para direções diferentes, porém com patamares diferenciados. Já a mista representa de um caso em que os variogramas diferem tanto na amplitude, quanto no patamar. Os diferentes tipos de anisotropias podem ser observados na Figura 7.

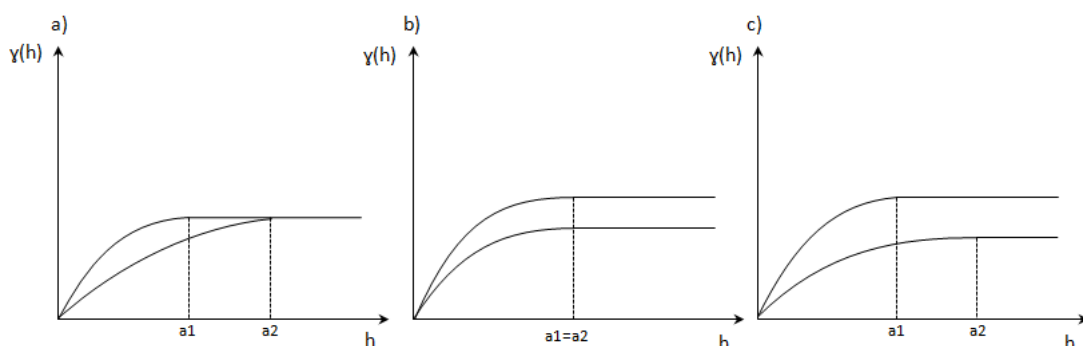


Figura 7: Variogramas de diferentes direções, sobrepostos de modo exemplificar respectivamente, da esquerda para a direita, uma anisotropia geométrica, uma anisotropia zonal e uma anisotropia mista (Yamamoto, 2001).

Conhecendo-se as propriedades teóricas dos variogramas, será possível ajustar modelos teóricos aos variogramas experimentais, facilitando a compreensão da variância espacial dos dados amostrados.

3.4 – O Modelo Multigaussiano

A distribuição Gaussiana é muito aplicada em estudos estatísticos, por ser muito conveniente para o cálculo de estimativas. A distribuição normal, ou Gaussiana, apresenta a maioria dos dados em torno da média (que também corresponde à mediana). A distribuição Gaussiana padrão apresenta média $m=0$ e variância $\sigma^2=1$ e, para qualquer variável aleatória Gaussiana, a função de densidade de probabilidade Gaussiana pode ser reescrita em função de m e σ^2 (Pilger, 2005).

Segundo Deutsch & Journel (1992), o modelo gaussiano de uma função aleatória é extraordinariamente conveniente, haja visto que qualquer função de distribuição cumulativa multivariada é completamente definida partindo-se do conhecimento da função de covariância.

A hipótese fundamental do modelo multiGaussiano assume que os vetores dimensionais da variável aleatória em estudo apresentam o comportamento de uma distribuição multivariada Gaussiana, e essa hipótese garante que a distribuição condicional da própria variável aleatória seja Gaussiana (Chilès e Delfiner, 1999 e Pilger, 2005).

De acordo com Deutsch e Journel (1992) e Batiston (2010), caso a função aleatória que define $Z(\mathbf{u})$ seja Gaussiana multivariada, a estimativa por krigagem simples (k_s), bem como sua variância definirão a média e a variância da dclp (distribuição condicional local de probabilidade cumulativa).

Tal estimativa e variância de $Z(\mathbf{u})$, como visto em Deutsch e Journel (1992) são obtidas partindo-se da krigagem simples realizada para um ponto não amostrado $z(\mathbf{u})$, com os dados dos n pontos amostrados relevantes ao redor, seguindo-se basicamente a seguinte equação:

$$Z_{ks}^*(\mathbf{u}) = \sum_{\alpha=1}^n \lambda_{\alpha}(\mathbf{u}) Z(\mathbf{u}_{\alpha}) + \left[1 - \sum_{\alpha=1}^n \lambda_{\alpha}(\mathbf{u}) \right] m$$

Sendo $Z(\mathbf{u})$ o modelo de variável aleatória estacionário, com média m e covariância $C(\mathbf{h})$. Os pesos $\lambda_{\alpha}(\mathbf{u})$ são determinados para minimizar-se a variância de estimativa, que resulta na equação:

$$\sum_{\beta=1}^n \lambda_{\beta}(\mathbf{u}) C(\mathbf{u}_{\beta} - \mathbf{u}_{\alpha}) = C(\mathbf{u} - \mathbf{u}_{\alpha}), \forall \alpha = 1, \dots, n$$

A variância de estimativa minimizada correspondente, também denominada variância de krigagem pode então ser obtida por:

$$\sigma_{k^2}^2(\mathbf{u}) = C(\mathbf{0}) - \sum_{\alpha=1}^n \lambda_{\alpha}(\mathbf{u}) C(\mathbf{u} - \mathbf{u}_{\alpha}) \geq 0$$

Contudo, para que tal aproximação possa ser aplicada, é necessário que a variável aleatória que define $Z(\mathbf{u})$ tenha comportamento multivariado Gaussiano e, muitas vezes, o histograma de dados não apresenta comportamento normal. Nesses casos, de acordo com Deutsch e Journel (1992), em geral, aplica-se a transformação gaussiana, sendo então adotado o modelo multiGaussiano $Y(\mathbf{u})$ para os dados transformados. Os processos de krigagem e/ou simulação poderão então ser executados sobre os y-dados obtidos através da referida transformação, tendo seus resultados propriamente revertidos a valores z pela transformação inversa.

3.5 – Transformação Gaussiana

Para que as técnicas de simulação ditas Gaussianas sejam passíveis de aplicação, é necessário que os dados analisados apresentem distribuição normal (Gaussiana), porém esta característica não é observada na maioria das variáveis. Para lidar com os dados muito discrepantes, Goovaerts (1997) sugere, entre outras possibilidades, a transformação dos dados, utilizada para se diminuir a influência de outliers, sem eliminá-los do conjunto de dados. Deve ter-se em mente, porém que transformar os dados implica em algumas dificuldades extras, como a transformação de volta dos resultados obtidos, que tende a exagerar qualquer erro associado com os processos de interpolação. Uma vez que a distribuição dos dados amostrados não é normal, opta-se pela transformação destes para a obtenção de uma nova distribuição, desta vez, gaussiana. Apesar do problema da transformação inversa dos dados, o método da simulação estocástica muito provavelmente ainda será altamente confiável.

Deutsch (2002) destaca que mesmo não havendo necessidade estrita dos dados apresentarem distribuição gaussiana, outros tipos de distribuição levariam a uma distribuição incorreta dos dados simulados.

De acordo com Olea (1999), se os valores das amostras não apresentam uma distribuição normal, tal característica pode ser alcançada através de uma transformação não linear que pode ser a Anamorfose Gaussiana ou a *normal score transformation*.

Deutsch (2002) atenta, porém, ao fato de que as técnicas que utilizam a distribuição gaussiana apresentam máxima entropia, o que implica em uma desordem espacial máxima. No caso, tal característica implica que os resultados obtidos através da simulação gaussiana não

estariam espacialmente estruturados tais quais os dados originais. Seria improvável, por exemplo, simular algum tipo de fluxo para se obter uma “imagem perfeita”.

De acordo com Pilger (2005), a transformação consiste basicamente de tomar-se a distribuição original $z(\mathbf{u}_\alpha)$ ($\alpha = 1, \dots, n$) e moldá-la em uma nova distribuição dos n valores, $y(\mathbf{u}_\alpha)$ ($\alpha = 1, \dots, n$), utilizando-se da função anamorfose Gaussiana, usualmente referida como $\phi(\cdot)$, conforme:

$$z(\mathbf{u}_\alpha) = \phi(y(\mathbf{u}_\alpha))$$

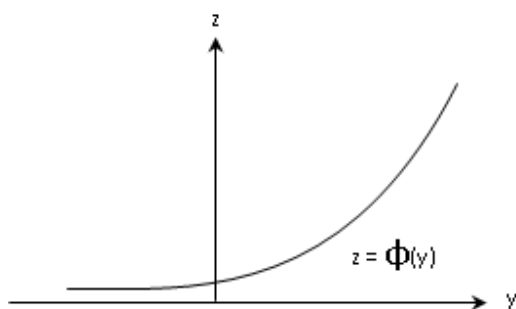


Figura 8: Função anamorfose Gaussiana

A partir de então o modelo multigaussiano poderia ser aplicado, sobre os valores transformados, dispostos na distribuição $y(\mathbf{u}_\alpha)$ ($\alpha = 1, \dots, n$). Vale ressaltar que, de acordo com Pilger (2005) e Batiston (2010), pelo fato de a função anamorfose gaussiana ser bijetora (monotônica crescente) a dclp da variável original pode ser diretamente acessada, pois as FAs (Funções Aleatórias) $F(z)$ e $G(y)$, referentes respectivamente às VAs (Variáveis Aleatórias) $Z(\mathbf{u})$ e $Y(\mathbf{u})$, apresentam a mesma probabilidade cumulativa para os mesmos quantis, como pode ser observado na Figura 9.

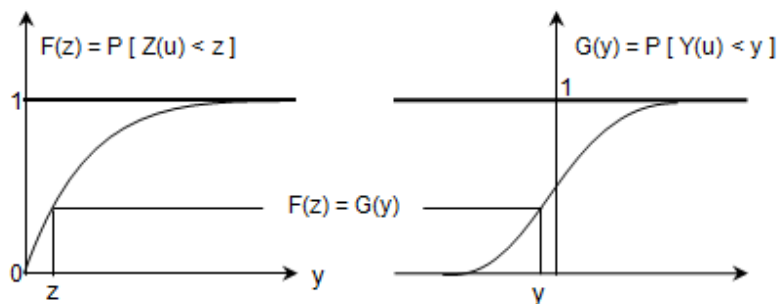


Figura 9: Comparação das Funções Aleatória $F(z)$ e $G(y)$, Riovardi (1991).

3.6 – Verificação da Normalidade Bivariada

Antes de se proceder à determinação da dclp, porém é necessária a averiguação da multigaussianidade dos dados transformados, ou seja, que a distribuição constituída pelos n valores da FA $y(\mathbf{u}_\alpha)$ ($\alpha = 1, \dots, n$) apresente um comportamento multivariado gaussiano para qualquer par de valores $Y(\mathbf{u})$ e $Y(\mathbf{u}+\mathbf{h})$ para todo \mathbf{u} e \mathbf{h} (Deutsch & Journel, 1992 e Pilger, 2005).

De acordo com Pilger (2005), apenas a normalização da distribuição original não garante que os dados transformados sigam tal comportamento, uma vez que este procedimento deve garantir apenas a normalidade univariada, carecendo de averiguação da multigaussianidade bivariada, trivariada, ..., n -variada. Tal comprovação, porém, trata-se em geral de um processo extremamente complicado, de modo a ser usualmente suficiente comprovar apenas a bigaussianidade.

Deutsch e Journel (1992) apresentam uma forma de se comprovar tal característica, que consiste da comparação de covariogramas experimentais dos indicadores (dos dados originais) e covariogramas teóricos para quantis definidos. Para tanto deve ser analisado se a função de distribuição cumulativa experimental de quaisquer pares de dados $\{y(\mathbf{u}_\alpha), y(\mathbf{u}_\alpha + \mathbf{h}), \alpha = 1, \dots, N(\mathbf{h})\}$, seja normal bivariada padrão para com uma função de covariância $C_Y(\mathbf{h})$, que através de relações analíticas tabeladas pode ser associada a qualquer função de distribuição cumulativa normal padrão da seguinte forma:

$$Prob \{Y(\mathbf{u}) \leq y_p, Y(\mathbf{u} + \mathbf{h}) \leq y_p\} = p^2 + \frac{1}{2} \int_0^{\arcsin C_Y(\mathbf{h})} \exp\left(-\frac{y_p^2}{1 + \sin\theta}\right) d\theta$$

sendo que $y_p = G^{-1}(p)$ é o p -quantil normal padrão e $C_Y(\mathbf{h})$ é o correlograma da função aleatória padrão $Y(\mathbf{u})$.

Então a probabilidade bivariada vinda da equação anterior é a covariância não-centrada do indicador para o limiar y_p :

$$Prob \{Y(\mathbf{u}) \leq y_p, Y(\mathbf{u} + \mathbf{h}) \leq y_p\} = E\{I(\mathbf{u}; p) \cdot I(\mathbf{u} + \mathbf{h}; p)\} = p - \gamma_I(\mathbf{h}; p)$$

sendo $I(\mathbf{u}; p) = 1$, se $Y(\mathbf{u}) \leq y_p$; e zero, caso contrário; e $\gamma_I(\mathbf{h}; p)$ é o semivariograma dos indicadores para a limiar do p -quantil y_p .

A checagem consistirá então em comparar o semivariograma dos indicadores dos dados amostrados à sua expressão normal bivariada. Alguns programas permitem que tal comparação seja feita, dispondo diretamente os semivariogramas dos indicadores sobre a binormal e, neste caso, utilizou-se da ferramenta “bigauss” do Software GeoVisual, disponível no Laboratório de Informática Geológica do Instituto de Geociências da Universidade de São Paulo.

3.7 – Simulação Estocástica

Segundo Olea (1999), ao se observar as estimativas realizadas através da krigagem, é possível perceber o efeito de suavização, ou seja, valores não amostrados que poderiam ser muito elevados foram subestimados, e valores anormalmente baixos foram superestimados. Isto ocorre devido ao fato de as interpolações feitas por krigagem suavizarem a variância, e, consequentemente o variograma, o que implica na perda da precisão global. Olea (1999) então, aponta a simulação estocástica para se contrapor ao efeito de suavização.

Deutsch (2002), afirma que tal fator de suavização se dá devido ao fato de a variância da estimativa ser calculada combinando-se somas lineares de n dados com seus devidos pesos na equação. Como em uma distribuição normal temos um volume maior de dados ao redor da média, as estimativas tendem a se distanciar dos valores máximo e mínimo.

Segundo Deutsch (2002), o efeito de suavização é diretamente proporcional à variância de krigagem, de modo a não haver suavização onde a variância de krigagem for igual a zero, bem como à estimativa estar totalmente suavizada longe de pontos amostrados, uma vez que todas as estimativas seriam iguais à média, e a variância de krigagem seria a variância estacionária $C(0)$. Ainda de acordo com Deutsch (2002), o efeito de suavização pode ser calculado da seguinte forma:

$$\text{Variância do Erro} = C(0) - \sum_{\alpha=1}^n \lambda_{\alpha} C(u, u_{\alpha})$$

Apesar de ser possível calcular a variância da estimativa feita pela krigagem (ou seja, existe uma forma de quantificar quão suavizadas estariam as estimativas) como visto em Deutsch (2002), Goovaerts (1997) ressalta que mapas suavizados não devem ser empregados em situações em que se torna necessário atentar para a presença de valores extremos. Como exemplo, o autor afirma que em um mapa de fluxo de água subterrânea, se suavizado, possivelmente se perderiam dados sobre locais onde a passagem da água seria facilitada ou dificultada, o que pode culminar no cálculo errôneo do tempo de viagem da água dentro do aquífero. Similarmente, o risco de poluição do solo por metais pesados poderia ser subestimado

em mapas krigados que não permitissem a presença de concentrações estimadas acima de um máximo tolerável.

A simulação estocástica é definida por Deutsch e Journel (1992) como o processo de se extrair realizações conjuntas, alternativas e equiprováveis de uma variável aleatória. Cada imagem gerada para cada uma das realizações é denominada “imagem estocástica”, e refletem as propriedades que foram impostas no modelo de função aleatória $Z(x)$. Assim sendo, quanto mais propriedades forem inferidas do conjunto amostral e incorporados no modelo $Z(x)$, melhor o modelo resultante. Deutsch & Journel (1992) ressaltam ainda que a simulação estocástica difere dos algoritmos de interpolação principalmente por valorizarem a precisão global acima da local, além de apresentar inúmeras representações de modelos numéricos alternativos igualmente prováveis para se atribuir um valor a pontos não amostrados.

Este tipo de ferramenta visa a obtenção de resultados não determinísticos, ou seja, existe mais de um resultado possível para cada ponto em que o algoritmo foi executado. O termo estocástico é atribuído devido a esta característica, e segundo Caers (2000), uma vez que a simulação estocástica reproduz a variabilidade espacial, suas estatísticas, seu variograma e o histograma, serão idênticos aos do espaço amostral.

Olea (1999) ressalta que a simulação estocástica não é uma solução definitiva, uma vez que em qualquer realização os erros tendem a ser muito maiores que em estimativas por krigagem. De modo que ao se optar por um método, deve-se ter em mente o que será mais importante, a continuidade espacial dos dados (fornecida pelos métodos de simulação) ou menores erros em estimativas locais (fornecida pelos métodos de krigagem).

De acordo com Srivastava (1994), apesar de técnicas estocásticas poderem gerar diversos resultados, muitos métodos aproveitam apenas um para a predição de dados não amostrados. Isto não se dá devido aos resultados serem ou não plausíveis, mas de reproduzirem uma heterogeneidade realista. As predições sobre os dados analisados, inclusive, são mais precisas quando baseadas em modelos que reflitam a heterogeneidade dos próprios dados.

Existem diversos métodos estocásticos de simulação, como a simulação sequencial Gaussiana (empregada neste trabalho de formatura), a simulação por Bandas Rotativas, em que as variáveis são simuladas em linhas unidimensionais que combinadas formam um modelo bi ou tridimensional, a decomposição LU (Lower/Upper) que envolve o uso de uma matriz $N \times N$, em que o número de pontos amostrados N pode chegar à casa dos milhões em alguns reservatórios, entre outros (Deutsch, 2002).

3.7.1 – A Simulação Sequencial Gaussiana

Os métodos de simulação sequencial em geral permitem, através de uma aproximação, gerar um valor para uma variável $Z(\mathbf{u})$ de sua distribuição condicional, sendo este influenciado por todos os dados disponíveis na vizinhança de \mathbf{u} , o que inclui os pontos amostrados e os previamente simulados (Deutsch e Journel, 1992).

De acordo com Srivastava (1994), os métodos de simulação sequencial, no entanto divergem no que se refere à dclp. A krigagem Multi-Gaussiana, por exemplo, gera suas estimativas baseada na ideia de que os valores amostrados seguem uma distribuição normal, ou gaussiana, de modo a ter em pauta sua média e desvio padrão. Este algoritmo, quando utilizado em um processo de simulação sequencial, geralmente é referido como simulação sequencial Gaussiana.

A variável será simulada sequencialmente, de acordo com a sua dclp normal caracterizada através da krigagem simples. A simulação condicional de uma variável contínua segundo Deutsch e Journel (1992) apresenta o algoritmo:

- Determina-se a função de distribuição condicional $F_z(z)$ representativa de toda a área de estudo, não apenas dos dados amostrados z disponíveis. Métodos de desagrupamento podem ser empregados, caso os dados estejam agrupados de em alguns pontos.

- Utilizando a função de distribuição cumulativa $F_z(z)$, realizar a transformação Gaussiana dos dados z em dados y que constituam uma função de distribuição cumulativa normal padrão.

- Checar ao menos a normalidade bivariada dos dados normalizados, para que se possa aplicar o modelo Gaussiano multivariado.

- Se o modelo de função aleatória Gaussiana multivariada puder ser aplicado para as variáveis y , é possível proceder à simulação sequencial Gaussiana da seguinte forma:

1. Definir um caminho aleatório que visite uma vez cada ponto da grade. A cada ponto \mathbf{u} , obter um número específico de dados vizinhos condicionantes, o que incluirá tanto os pontos amostrais quanto os previamente simulados;

2. Usar da krigagem simples com o variograma feito a partir dos dados normalizados para determinar os parâmetros (média e variância) da dclp da variável aleatória $Y(\mathbf{u})$ na localização \mathbf{u} ;
3. Escolher um valor simulado $y^{(j)}(\mathbf{u})$ desta dclp;
4. Adicionar este valor $y^{(j)}(\mathbf{u})$ escolhido ao banco de dados. Os passos de 1 a 4 podem ser observados na Figura 10;

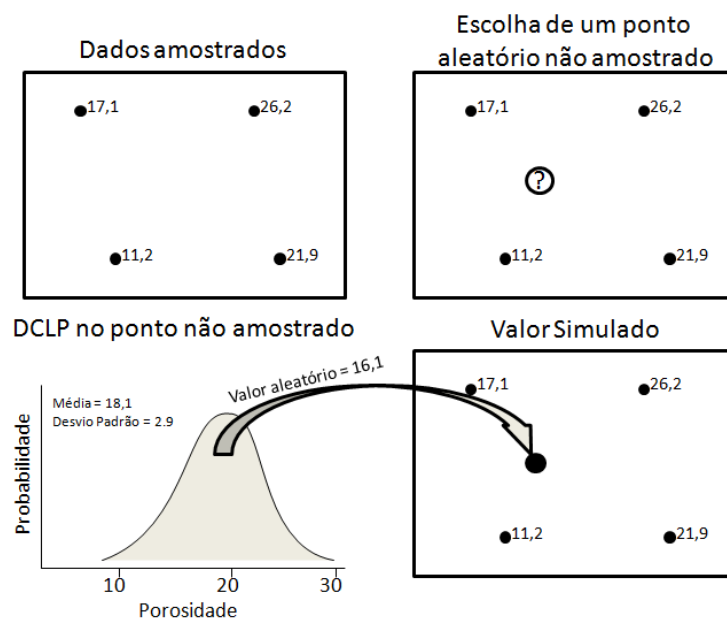


Figura 10: Passos do algoritmo da ssG
(Srivastava, 1994)

5. Proceder ao próximo ponto, reiniciando o processo, até que todos os pontos sejam simulados.

-Executar a transformação inversa dos dados normais simulados $\{y^{(j)}(\mathbf{u}), \mathbf{u} \in A\}$ em valores simulados da variável original $\{z^{(j)}(\mathbf{u}) = \varphi^{-1}(y^{(j)}(\mathbf{u}), \mathbf{u} \in A)\}$

Basicamente, portanto, a simulação sequencial Gaussiana se trata do processo de gerar realizações parciais utilizando de funções aleatórias normais multivariadas para a obtenção de valores para pontos não amostrados (Olea, 1999 e Furuie 2009). Isso consistirá da atribuição de valores através da determinação da dclp em cada um dos pontos não amostrados que se deseja simular (Caers, 2000)

4- Tratamento dos dados

Para o desenvolvimento deste trabalho foram utilizados os programas Isatis e Geovisual, disponíveis no Laboratório de Informática Geológica (LIG) do Instituto de Geociências da Universidade de São Paulo.

4.1 – Análise estatística dos dados

Na Tabela 1 as apresentadas as estatísticas descritivas da variável de interesse e nas Figuras 11 e 12 pode-se observar, respectivamente, a curva de frequência acumulada e o histograma. O histograma indica uma distribuição assimétrica positiva para os teores de cádmio.

Tabela 1: Estatísticas descritivas da variável Cd.

Número de dados: 259
Média: 1,309
Desvio Padrão: 0,913
Coeficiente de Variação: 0,698
Valor Máximo: 5,129
Quartil Superior: 1,713
Mediana: 1,065
Quartil Inferior: 0,629
Valor Mínimo: 0,135

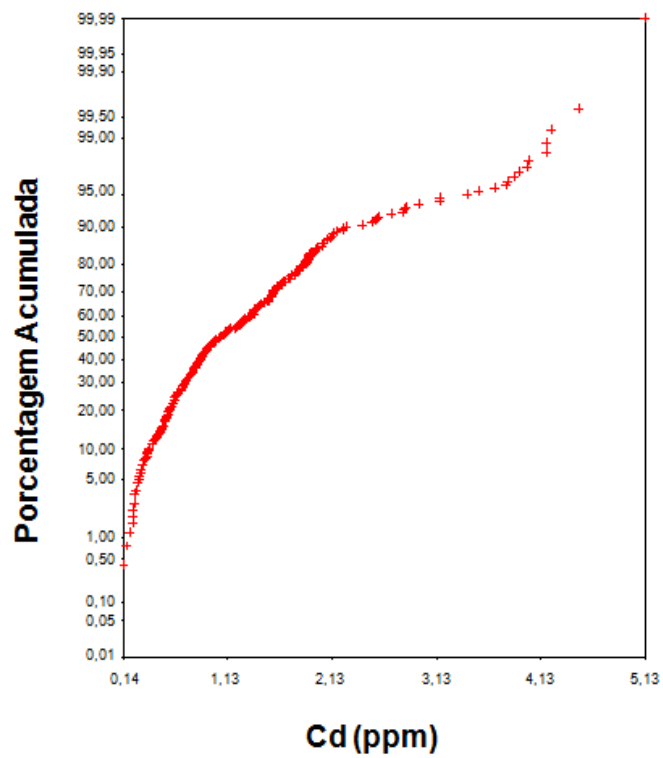


Figura 11: Curva cumulativa dos valores de cádmio dos dados amostrados.

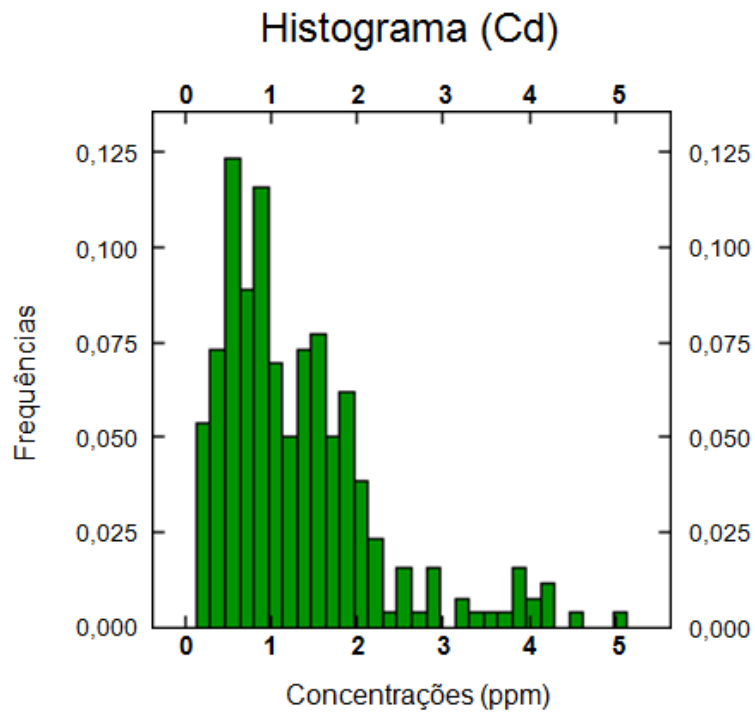


Figura 12: Histograma dos teores de cádmio (ppm)

Tratando-se de um caso de possível contaminação, os valores muito elevados foram analisados em detalhe, observando-se como se distribuem espacialmente, bem como a proximidade a outros pontos de concentrações também elevadas.

As amostras que apresentaram teores elevados foram realçadas na Figura 13A, e é possível observar que se distribuem no mapa aparentemente sem um padrão, como pode ser observado na Figura 13B, porém, visando analisar como se apresentam os valores inferiores adjacentes no histograma, estes foram demarcados na Figura 14A e 1A, gerando respectivamente os mapas nas Figuras 14B e 15B.

Comparando-se os três mapas, nota-se que boa parte dos valores anômalos está relativamente concentrada em alguns pontos do mapa, grande parte, inclusive, nos arredores dos pontos de teores mais elevados (a partir de 3,8). Tal arranjo possibilita evidenciar as áreas com a possível liberação ou acúmulo dos contaminantes, que devem ser levadas em consideração no momento da simulação das localidades não amostradas.

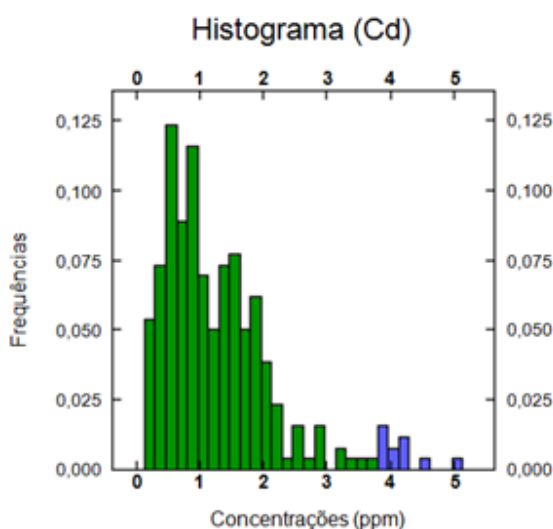


Figura 13A: Histograma com demarcação de valores muito elevados

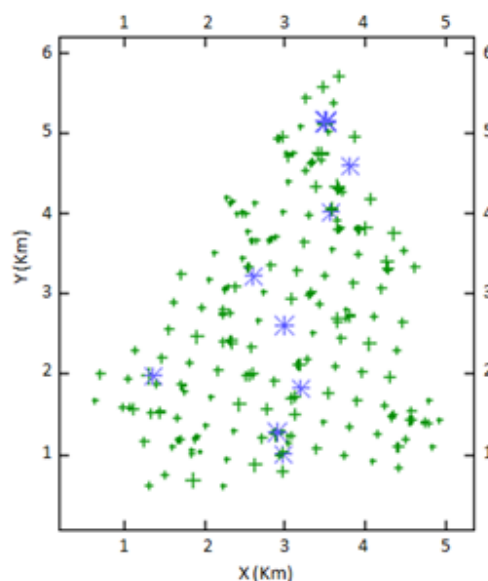


Figura 13B: Mapa com a localização dos valores demarcados na Figura 13A

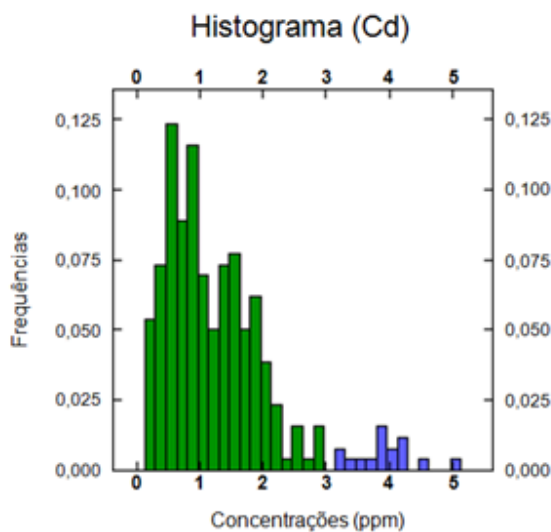


Figura 14A: Histograma com demarcação de valores muito elevados, com a adição das classes vizinhas

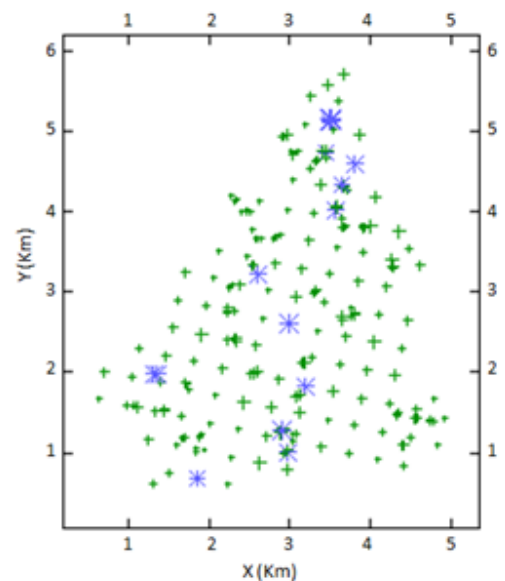


Figura 14B: Mapa com a localização dos valores demarcados na Figura 14A

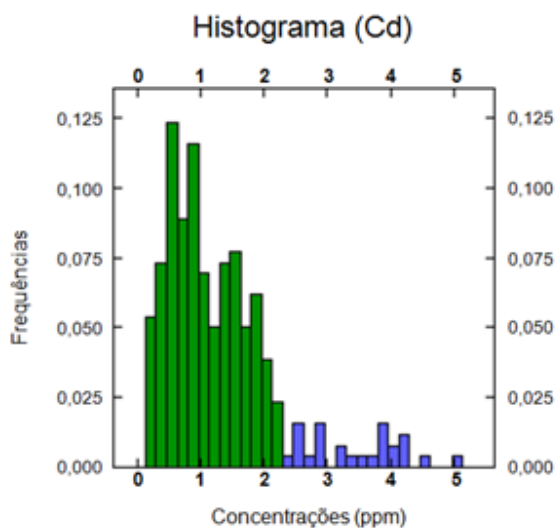


Figura 15A: Histograma com demarcação de valores elevados

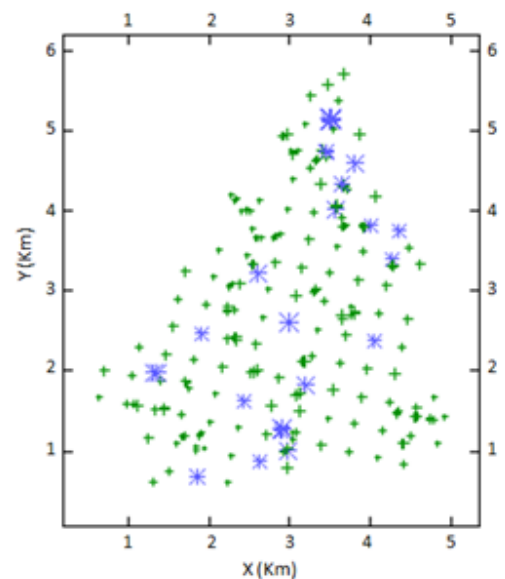


Figura 15B: Mapa com a localização dos valores demarcados na Figura 15A

4.2 – Análise geoestatística dos dados não transformados

A princípio gerou-se um variograma omnidirecional com modelo ajustado, (apresentado na Figura 16), de modo a analisar como os dados se comportam de forma global. É possível inferir a amplitude igual a 430 metros e patamar igual a 0,8. O variograma é descontínuo próximo à origem e apresentar efeito pepita com valor interpretado de 0,2.

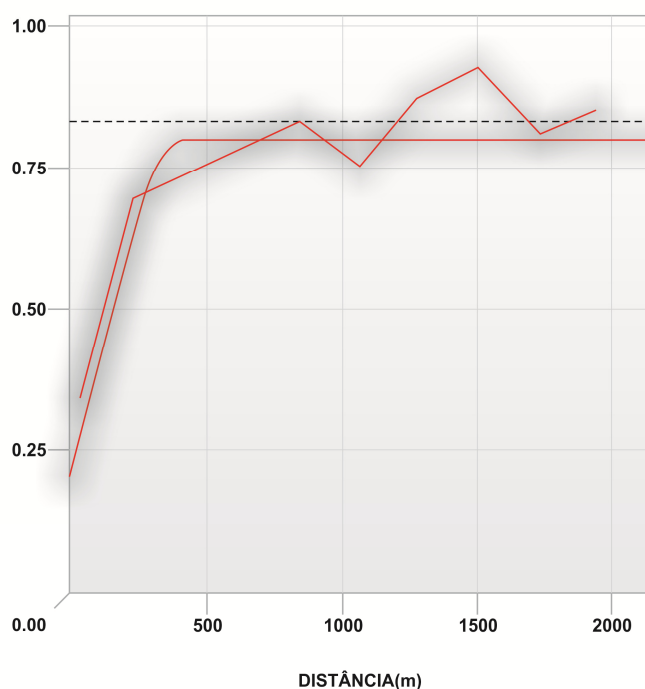


Tabela 2: Características do Variograma Omnidirecional

Tolerância Angular: 90°
Passo: 215
Efeito Pepita: 0,2
Amplitude: 430m
Patamar: 0,8

Figura 16: Variograma omnidirecional

Verificando-se se há anisotropia, foi feita a análise exploratória dos dados, calculando-se os variogramas para quatro direções, 0°, 45°, 90° e 135°, conforme pode ser observado na Figura 17.

Analizando-se o resultado, definiu-se que existe anisotropia geométrica e que a direção de 90° apresenta a maior amplitude, sendo, portanto a 0° como a de menor amplitude. A Figura 18 mostra o variograma experimental para as direções de 0 e 90°. Assim como no variograma omnidirecional, estimou-se o efeito pepita em 0,2

Para que fosse possível observar de melhor forma cada um dos variogramas, foram geradas as Figuras I.1 a I.4, apresentadas no Anexo I. As propriedades de cada variograma podem ser observadas na Tabela 3.

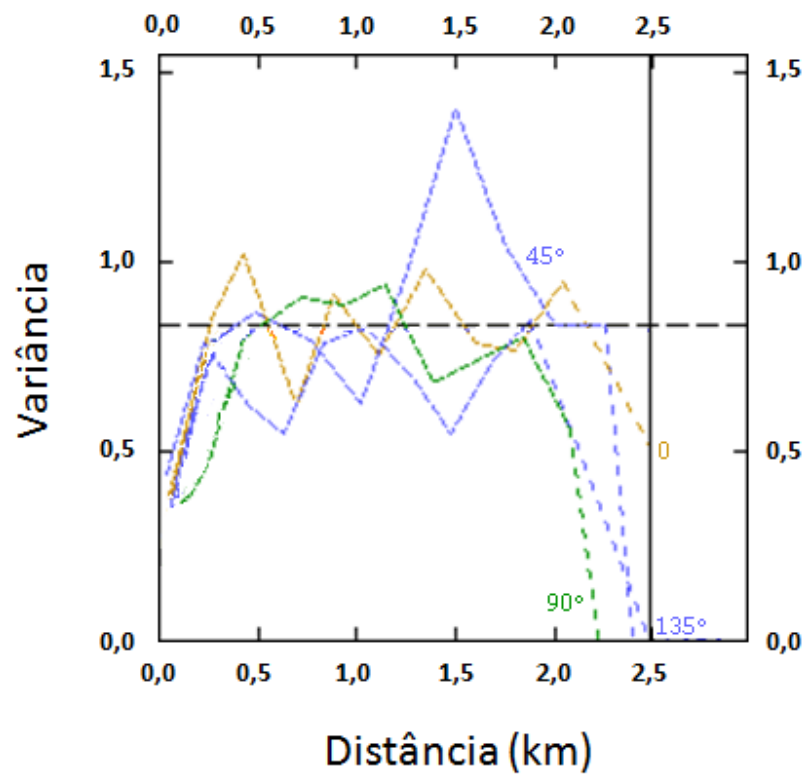


Figura 17: Variograma multidirecional, para 0°, 45°, 90° e 135°.

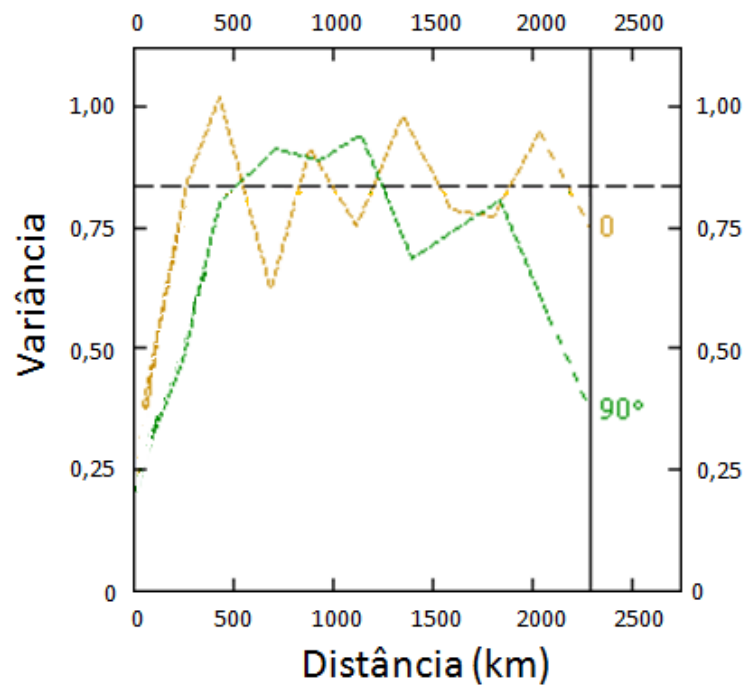


Figura 18: Variograma experimental para as direções 0° e 90°

Tabela 3: Características dos variogramas direcionais

Direção	0°	45°
Características	Tolerância angular: 22,5°	Tolerância angular: 22,5°
	Passo: 225	Passo: 250
	Efeito Pepita: 0,2	Efeito Pepita: 0,2
	Amplitude: 320m	Amplitude: 480m
	Patamar: 0,82	Patamar: 0,82m
Direção	90°	135°
Características	Tolerância angular: 22,5°	Tolerância angular: 22,5°
	Passo: 230	Passo: 210
	Efeito Pepita: 0,2	Efeito Pepita: 0,2
	Amplitude: 720m	Amplitude: 325m
	Patamar: 0,82	Patamar: 0,82

4.3 – Transformação dos dados

Para que os dados pudessem ser dispostos em uma distribuição normal, usou-se a função de anamorfose Gaussiana. O histograma dos resultados da transformação é apresentado na Figura 19, e as estatísticas descritivas da nova distribuição na Tabela 4.

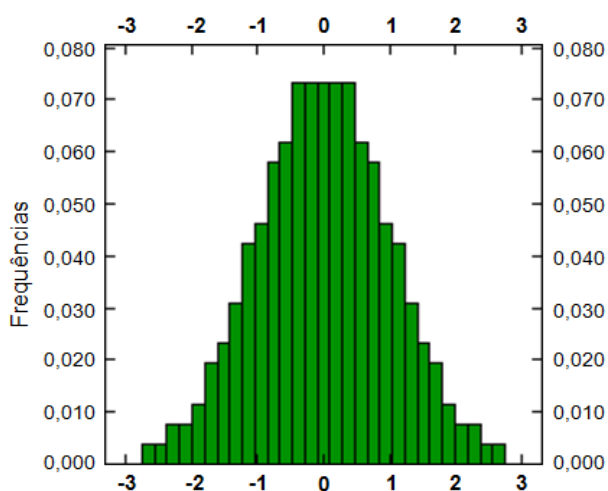


Figura 19: Histograma dos produtos da transformação

Tabela 4: Estatísticas

Básicas da distribuição dos dados transformados

Número de amostras: 259
Mínimo: -2,74
Máximo: 2,74
Média: 0
Desvio padrão: 0,99

Torna-se visível, agora, a diferença do novo histograma para com a distribuição original (Figura 20 e Figura 21, respectivamente, comparadas abaixo), obedecendo agora à distribuição gaussiana, ou seja, com grande parte dos valores distribuídos ao redor da média e com uma quantidade cada vez menor de valores conforme se aproxima às extremidades do histograma.

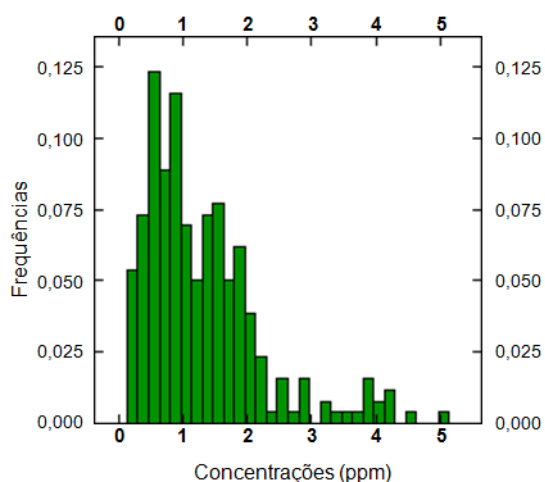


Figura 20: Histograma dos dados amostrados

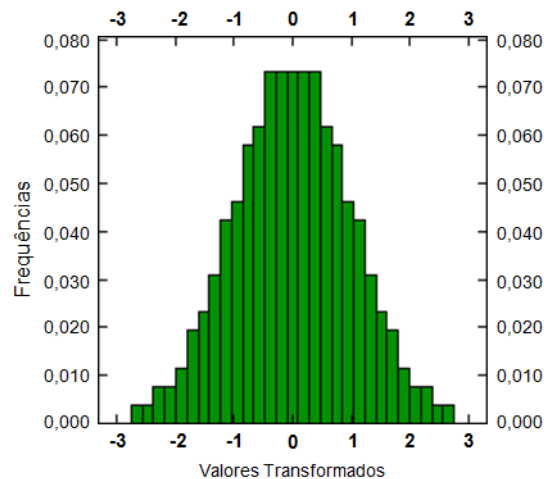


Figura 21: Histograma dos resultados da anamorfose

Para que a transformação dos dados possa ser aceita é necessário testar a biGaussianidade dos dados. Para tanto utilizou-se a ferramenta Bigauss do Geovisual. Os quartis escolhidos para o teste foram o 1º, o 2º e o 3º. Os resultados obtidos podem ser observados nas Figuras 22, 23 e 24, respectivamente

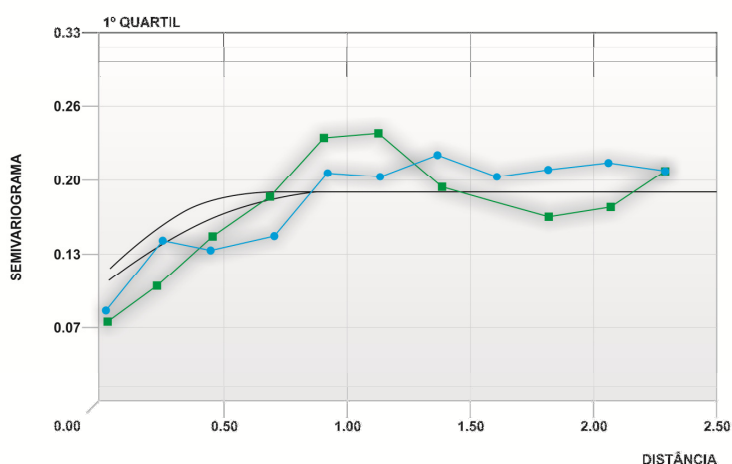


Figura 22: Semivariogramas dos indicadores para o 1º quartil

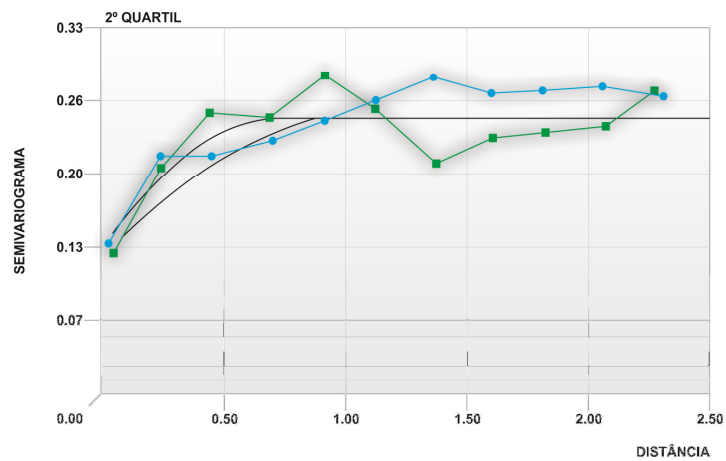


Figura 23: Semivariogramas dos indicadores para o 2º quartil

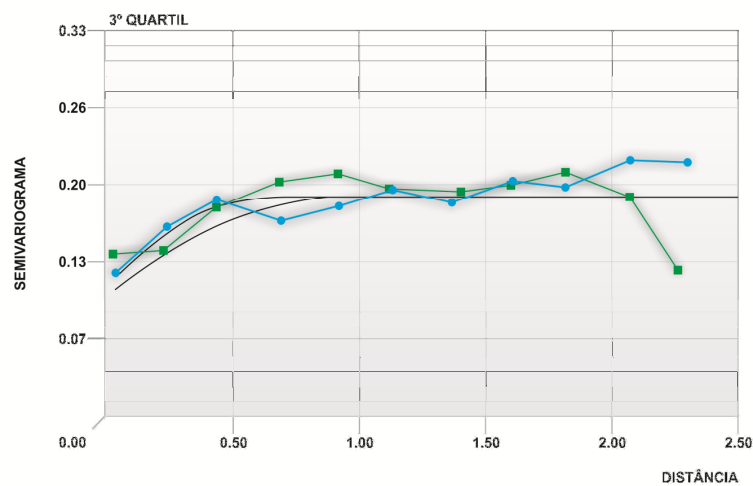


Figura 24: Semivariogramas dos indicadores para o 3º quartil

Os semivariogramas dos indicadores dos quartis dos dados comportam-se de forma próxima à normal bivariada para com as suas expressões teóricas, ou seja, comprova-se, portanto, o comportamento Gaussiano bivariado para os dados.

4.4 – Análise Geoestatística dos dados transformados.

Foram também calculados novos variogramas direcionais e omnidirecional, para que fosse possível observar a variância entre os resultados das transformações, e também comparar o comportamento dos novos dados com o dos dados amostrados. O novo variograma omnidirecional é apresentado na Figura 25.

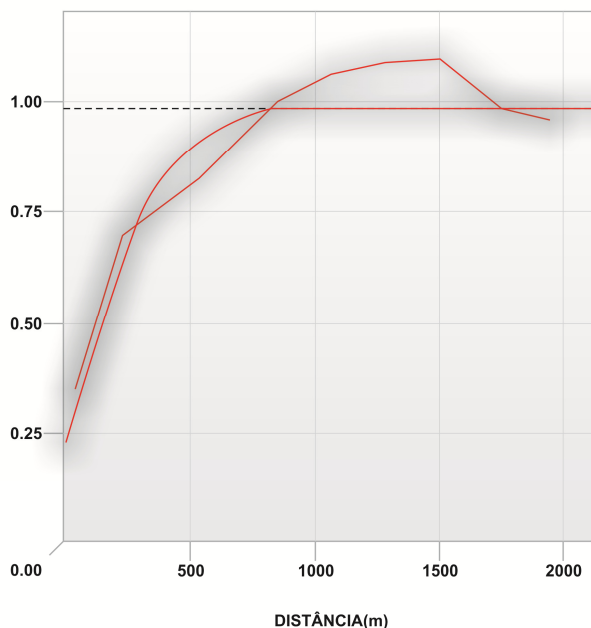


Figura 25: Variograma omnidirecional dos dados transformados

Tabela 5: Características do variograma omnidirecional dos resultados da anamorfose Gaussiana.

Tolerância Angular: 90°
Passo: 215
Efeito Pepita: 0,22
Amplitude das estruturas: 360m e 850m
Patamar: 0,98

Na Figura 26, observa-se o variograma experimental calculado para as direções de 0° e 90°, e o respectivo modelo ajustado. Pode-se notar que os dados transformados apresentam anisotropia, mista. A Tabela 6 apresenta os parâmetros do modelo ajustado.

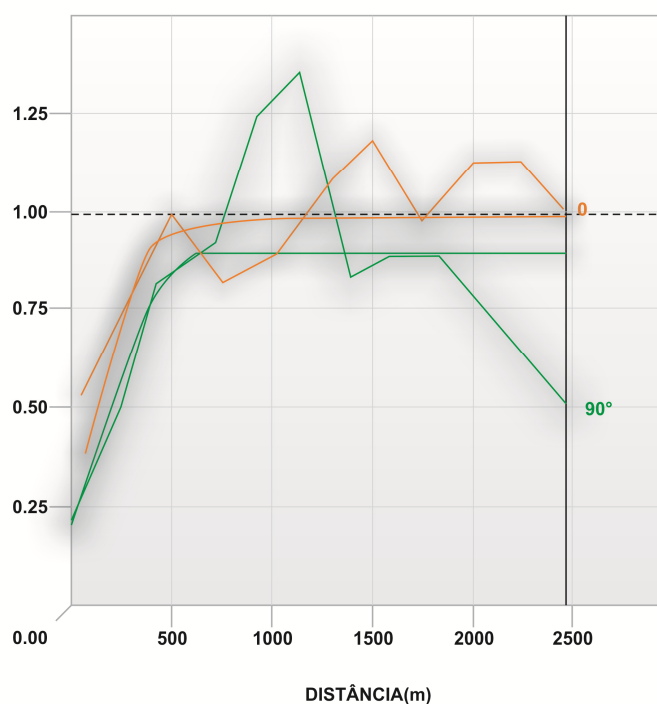


Tabela 6: Variogramas direcionais dos dados amostrados

0°	90°
Tolerância Angular: 22,5°	Tolerância Angular: 22,5°
Passo: 250m	Passo: 230m
Efeito Pepita: 0,2	Efeito Pepita: 0,2
Amplitudes das Estruturas: 434m e 1500m	Amplitude: 650m
Patamar: 0,98	Patamar: 0,895

Figura 26: Variogramas ajustados dos dados transformados para as direções de 0° e 90°.

Cabe ressaltar que foi feita a análise exploratória destes dados, porém como a elipse que representa a anisotropia não foi deslocada, optou-se por apresentar os variogramas calculados apenas ao longo dos eixos de maior e menor alongação.

Antes de proceder-se à simulação propriamente, é necessário que se comprove a normalidade multivariada dos dados, para que seja possível aplicar o modelo Gaussiano aos dados, possibilitando a obtenção da dclp para cada ponto a ser simulado.

4.5 – Definição do *Grid* de Simulação

De modo a definir os locais a serem simulados, foi construída a malha regular apresentada na Figura 27. Os parâmetros desta malha são apresentados na Tabela 7, apresentada abaixo.

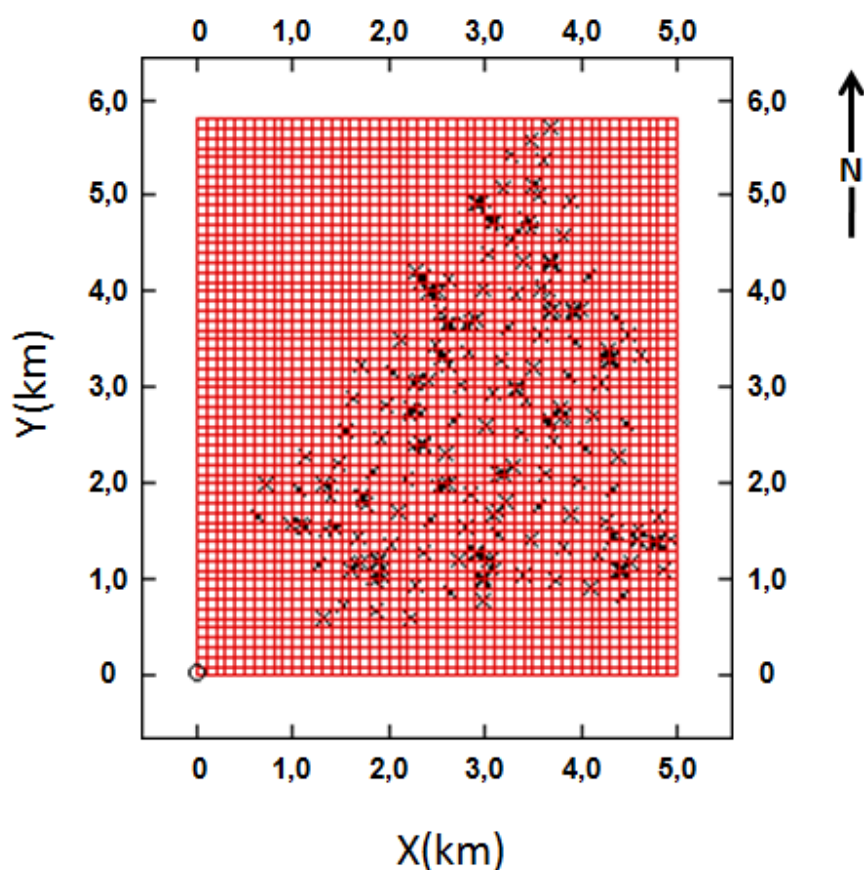


Figura 27: *Grid* de simulação

Tabela 7: Parâmetros do Modelo de Blocos

Área total simulada: 29 km ²
Área de cada bloco: 10.000 m ²
Número de blocos: 2900

4.6 – A aplicação da simulação sequencial Gaussiana

Foram feitas 500 imagens (realizações). Destas foram escolhidas 50 ao acaso, apresentadas em anexo, de modo a se demonstrar como se dá a distribuição dos possíveis valores nos locais não amostrados.

Devido às proporções da área abrangida pela amostragem, optou-se pela melhor resolução das imagens, dado que cada quadro do *grid* abrange uma área de aproximadamente 10.000 m². Em se tratando de casos onde uma possível intervenção será necessária para a

remediação de uma área contaminada, as dimensões de tal ação devem estar bem estabelecidas.

De modo geral, é possível perceber a tendência das áreas a noroeste e levemente a sul-sudeste do centro do *grid* (estendendo-se em algumas realizações para leste e oeste) de apresentarem as maiores incidências de concentrações elevadas de Cd (maiores que 3ppm de Cd em solo, em vermelho nas imagens das realizações). Menores teores de Cd (entre 0 e 0,5ppm, em tons azuis nas imagens das realizações), por sua vez, tenderão a apresentar-se levemente a norte e sul-sudoeste do centro do *grid*.

Em grande parte das realizações nota-se que a nordeste, noroeste e sudeste, áreas simuladas que encontram-se distantes dos pontos de amostragem, os resultados foram de valores basicamente médios (entre 1 e 2,5ppm), em comparação com a amostragem.

A partir das simulações, foram gerados mapas de probabilidade, que consistem de demonstrar geograficamente a probabilidade das concentrações do elemento de interesse superarem um teor pré-estabelecido.

4.6.1 – Pós-simulação

De acordo com a OMS (2000), cereais como o trigo e o arroz e outros vegetais, como a batata (que inclusive constituem a base da alimentação na maioria das culturas) tem uma grande capacidade de acumular cádmio (entre 10 e 300 µg/kg, a maior dentre os alimentos, a não ser por órgãos internos de alguns mamíferos).

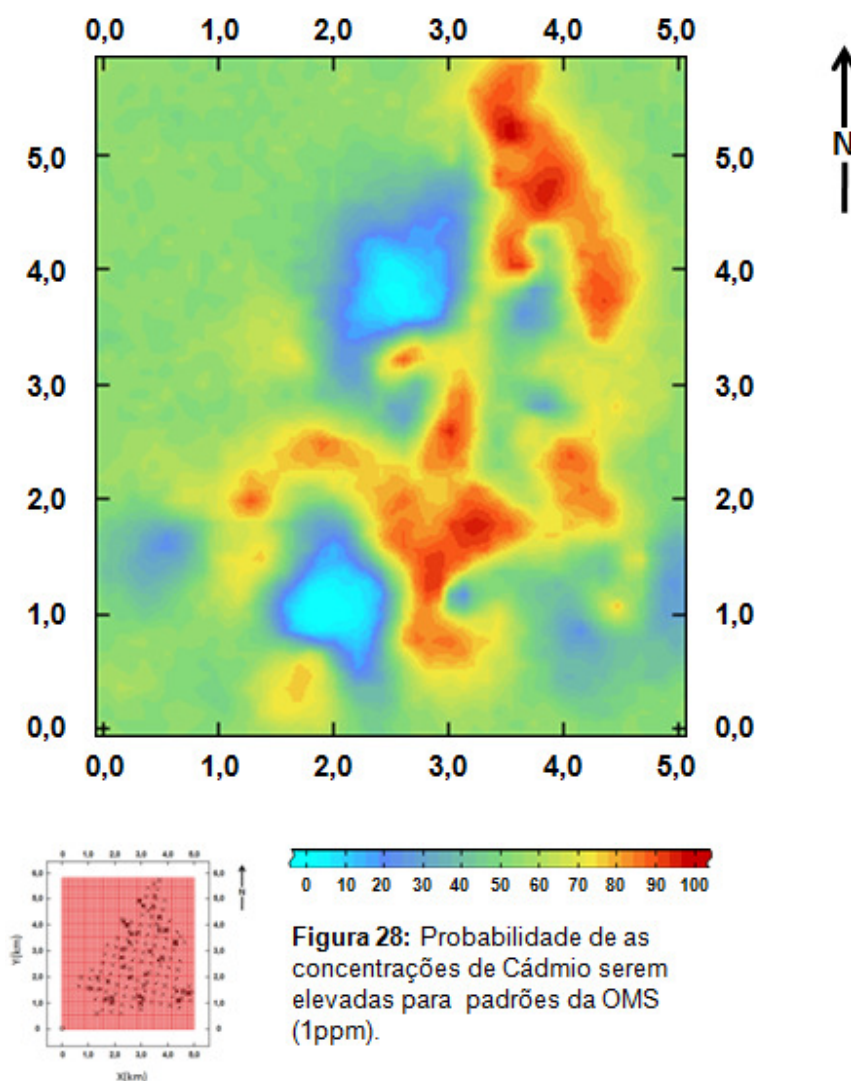
Fujimoto & Uchida (1978) ressaltaram a capacidade do arroz de retirar o cádmio diretamente do solo, não envolvendo necessariamente a passividade do vegetal durante absorção de água rica em cádmio. Plantações de arroz em áreas de solo contaminado por cádmio podem apresentar inclusive concentrações acima de 1000 µg/kg (OMS, 2000)

Para uma melhor compreensão sobre o significado dos dados simulados, foi gerado um mapa de probabilidade da concentração de cádmio no solo superar 1ppm, pois, mesmo que o cádmio não seja absorvido pelas plantações em totalidade, tal concentração já poderia gerar teores anômalos em vegetais como o trigo e a batata, que constituem a base da alimentação local. O mapa de probabilidades é apresentado na Figura 28.

É possível perceber que a possibilidade de as concentrações de cádmio ultrapassarem o valor de 1ppm é maior nas áreas citadas anteriormente como as de maior incidência de altas

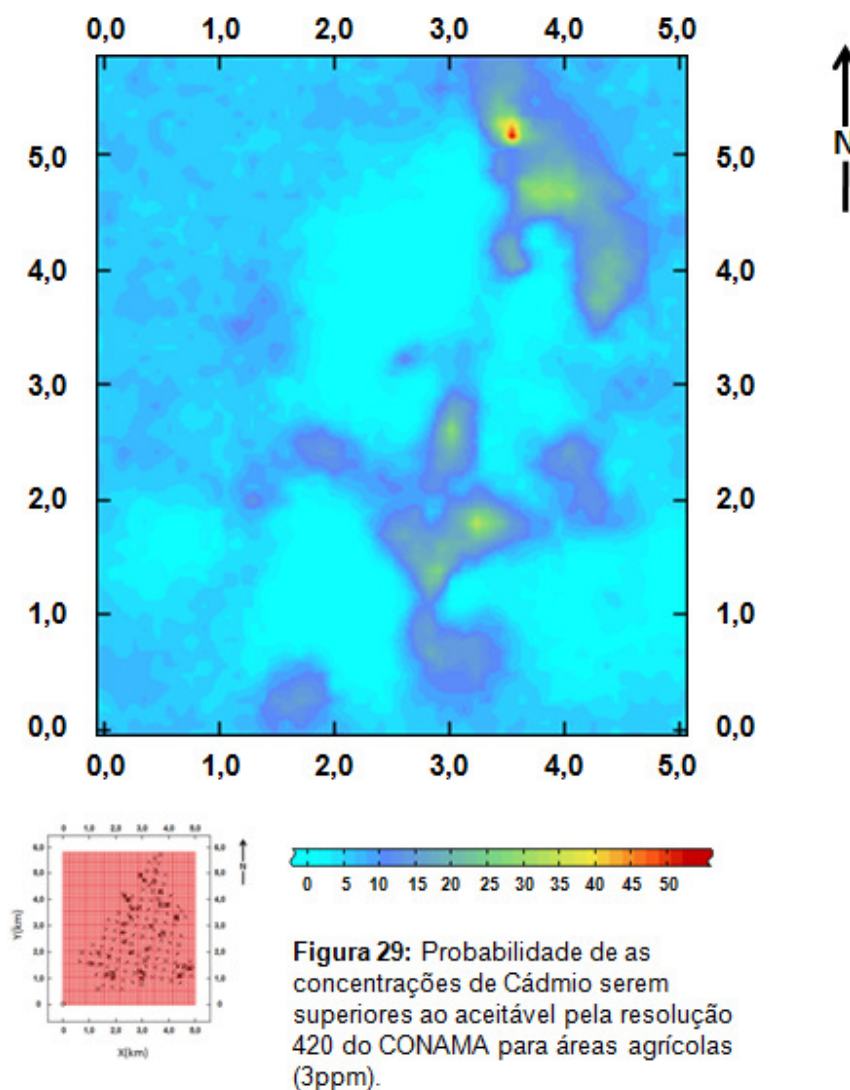
concentrações, de modo que o uso de tais regiões para a agricultura não seria aconselhável, sem antes haver alguma intervenção para que os teores do elemento estudado fossem reduzidos.

Dada sua elevada toxicidade e tendência de se acumular nos rins, causando uma série de patologias relacionadas à excreção, além de possíveis problemas relacionados à estrutura óssea em casos de exposição a teores elevados, mesmo que em curtos períodos de tempo, recomenda-se que áreas com histórico de contaminação por cádmio, principalmente por atividades industriais, não sejam utilizadas para agricultura.



No Brasil, porém, de acordo com a resolução 420 do Conselho Nacional do Meio Ambiente (CONAMA, 2009), para áreas agriculturáveis, são permitidos a presença de teores de cádmio em solo até 3ppm. Foi gerado então um segundo mapa de probabilidades (Figura 29).

É possível observar na Figura 29 que apenas uma região relativamente pequena a norte-nordeste teria uma probabilidade maior que 50% de apresentar teores suficientes para haver necessidade de intervenção, de acordo com os padrões do CONAMA (seguidos pela CETESB, no estado São Paulo).



Na Tabela 8 são comparadas algumas concentrações de cádmio aceitáveis de acordo com a EPA, órgão de proteção ambiental norte-americano, adotado em diversos estudos de

impacto ambiental ao redor do mundo, bem como os valores estipulados pelo CONAMA. Como os valores aceitáveis para áreas residenciais e industriais de ambas as listas superam em muito o valor máximo de concentração de cádmio nas amostras dos dados estudados, não são apresentados mapas de probabilidade de superação de tais valores, uma vez que a referida probabilidade não passaria de 0% em nenhuma localidade. Além disso, o Cd em solos não agrícolas em geral pouco afetará o homem, dado que o transporte do elemento em solo para o ser humano será feito através dos alimentos.

Vale ressaltar que, como visto em Atteia *et al.*(1994), os teores de cádmio em superfície dificilmente teriam qualquer relação com a geologia local, e como Jura é uma região de atividade principalmente agrícola, é recomendável o estudo pormenorizado nas localidades de maior probabilidade de ocorrência de valores elevados de cádmio, dependendo do tipo de cultivo a ser instalado, que pode vir a absorver quantidades não naturais do elemento.

Tabela 8: Valores de intervenção para Cd em solo

Norma	Finalidade do uso do solo na região		
	Agrícola	Residencial	Industrial
EPA	-	70ppm	800ppm
CONAMA 420	3ppm	8ppm	20ppm

Recomenda-se também a diminuição ou suspensão, se possível, de qualquer atividade que libere o contaminante, no caso muito provavelmente se tratando do uso de fertilizantes (principalmente os fabricados a base de fosfatos), além do controle na liberação de esgoto, fatores que liberam e influem na mobilidade do Cd em solo.

5- Conclusões

Não houve grandes dificuldades na utilização dos dados, oferecidos em Goovaerts (1997). Ao se comparar o histograma com o mapa de pontos amostrados demarcando-se os grupos de maior valor, foi possível observar que os pontos de maior concentração encontram-se relativamente próximos, agrupados ao longo do mapa, sem um padrão aparente.

Analisando as características geoestatísticas dos dados amostrados e dos dados transformados, foi possível observar que ocorre anisotropia. Orientou-se a elipse que representa esta anisotropia em 90° o eixo de maior alongação e em 0° o de menor.

Foram realizadas 500 simulações em 2D e as simulações demonstraram que as áreas de maior concentração de cádmio (em vermelho) provavelmente se localizariam a noroeste e levemente a sul-sudeste do centro do *grid*, por vezes alcançando regiões mais a leste ou oeste.

Comparando-se as análises realizadas a algumas normas atuais internacionais e nacionais, foi possível constatar que a concentração de cádmio na região de Jura é sim anômala, porém dificilmente tais concentrações representariam algum tipo de perigo real a curto prazo, a não ser em algumas localizações específicas. Grande parte do risco envolveria o tipo de plantio a ser exercido, bem como da ausência de qualquer intervenção para que o volume de cádmio no solo seja estabilizado ou reduzido.

As realizações da simulação sequencial Gaussiana, bem como a montagem do mapa de probabilidades, foram de suma importância para permitir a observação de áreas de maior periculosidade, ou que necessitarão de estudos específicos antes da instalação de qualquer cultura agrícola (áreas em vermelho da Figura 28).

Recomendar-se-ia para área, portanto, com base nas informações fornecidas pela OMS, o uso de ações remediadoras, ou ao menos de contenção do contaminante, haja visto que tais concentrações anômalas não são de origens naturais, sendo muito provavelmente resultado de atividades antrópicas, como o uso de fertilizantes a base de fosfatos e a liberação de esgoto não tratado.

6- Referências Bibliográficas

- Atteia O., Dubois J. P., Webster R., 1994, Geostatistical Analysis of Soil Contamination in the Swiss Jura. In: Environmental Pollution 86. p315-327.
- Batiston., E.L., 2010, Simulação Sequencial Gaussiana Usando Latin Hypercube Sampling: Estudo de Caso Minério de Ferro Carajás, Dissertação de Mestrado, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS. 88p
- Borrough, P. A., 1993. Soil Variability: A Late 20th Century View. In: Soils Fertil, 56. p529-562.
- Caers, J., 2000, Adding Local Accuracy to Direct Sequential Simulation, in Mathematical Geology, Vol. 32. n° 7. p815-850.
- CETESB, 2005 – Valores Orientadores para Solo e Águas Subterrâneas no Estado de São Paulo. 4p
- Chillés J. P. & Delfiner P., 1999, Geostatistics: Modeling Spatial Uncertainty, Wiley, New York.
- CONAMA, CONSELHO NACIONAL DO MEIO AMBIENTE 2009 – Resolução n° 420, de 28 de dezembro de 2009, do Conselho Nacional do Meio Ambiente. Brasília, 16 p.
- Deutsch C. V., & Journel A. G., 1992, GSLIB-Geostatistical Software Library and User's Guide. Oxford University, New York. 369p.
- Deutsch C. V., 2002, Geostatistical Reservoir Model. 376p
- Eggenberger, U. & Waber, H. N., 1998, "Cadmium in Seepage Waters of Landfills: A Statistical and Geochemical Evaluation, "Report of November 20, 1997 for the OECD Advisory Group on Risk Management Meeting".
- EPA, U.S. ENVIRONMENTAL PROTECTION AGENCY. 2009. Regional Screening Levels. Oak Ridge National Laboratory – ORNL. May, 2012
- Friberg, L., Piscator, M., Nordberg, M. B. 1971. Cadmium in the Environment. Ohio: C.R.C. Press, 166 p.
- Fujimoto, T. e Uchida, Y., 1978, Cadmium Absorption by Rice Plants, Soil Science and Plant Nutrition, 25:3, p407-415.

Furuie R. A., 2009, Estudo Comparativo de Métodos de Estimativas e Simulações condicionais, Universidade de São Paulo, Dissertação de Mestrado. 166p.

Goovaerts, P., 1997, Geostatistics for natural resources evaluation: Oxford University Press, New York, 483p.

Hogan, C.M., 2010, Toxicology: Heavy Metal, In: Encyclopedia of Earth. Eds. Cutler J. Cleveland (Washington, D.C.: Environmental Information Coalition, National Council for Science and the Environment). 9p

Journel, A.G., and Huijbregts, Ch. J., 1978. Mining Geostatistics. Academic Press (London), 600p.

Koch Jr., G. S. & Link, R. F., 1971. Statistical Analysis of Geological Data. New York Dover Publishing. Vol. I. 375p.

Kraft G. G. 1990. The Future of Cadmium Electroplating: Metal Finishing. [S.1]: Hackmack, jul. p 29-31;

NUS Corporation, 1987. Characterisation of Municipal Waste Combustor Ashes and Leachates from Municipal Solid Waste Landfills, Monofills, and Codisposal Sites, Report prepared for the U.S. Environmental Protection Agency, Office of Solid Waste, R-33-6-7-1, Washington, DC. 142p.

Olea R. A., 1999 Geostatistics for Engineers and Earth Scientists. 303p.

OMS, 2000, Air Quality Guidelines – Second Edition. WHO Regional Office for Europe, Copenhagen, Denmark. 228p

Pilger, G. G., 2005, Aumento da Eficiência dos Métodos Sequenciais de Simulação Condicional Tese de Doutorado, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS. 260p.

Riovard, J., 1991. Introduction au Krigage Disjonctif et à la Geostatistique Non Linéaire. Cours C-139, 2^{ème} edition, Centre de Geostatistique, EMP. 99p.

Srivastava, R. M., 1994, The Visualization of Spatial Uncertainty, Fss International, Vancouver, British Columbia, Canada. In: Yarus J. M. & Chambers R. L., Stochastic Modeling and Geostatistics, American Association of Petroleum Geologists. p339-345.

Tavares T. M., 1990. Avaliação de Efeitos das Emissões de Cádmio e Chumbo em Santo Amaro, Bahia, Tese de Doutorado, Universidade de São Paulo. 273p.

Webster R., Atteia O., Dubois J. P. ,1994, Corregionalization of Trace Metals in the Soil in the swiss Jura, In: European Journal of Soil Science 45, p205-218.

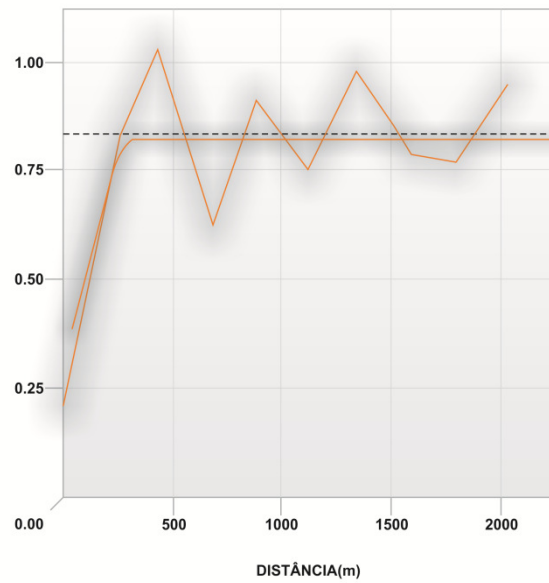
Weeks, M. E. & Leicester, H. M. 1968. Discovery of elements, 7.ed. Pennsylvania: Journal of Chemical Education, 896 p.

Yamamoto, J. K., 2001. Avaliação e Classificação de Reservas Minerais. 228p.

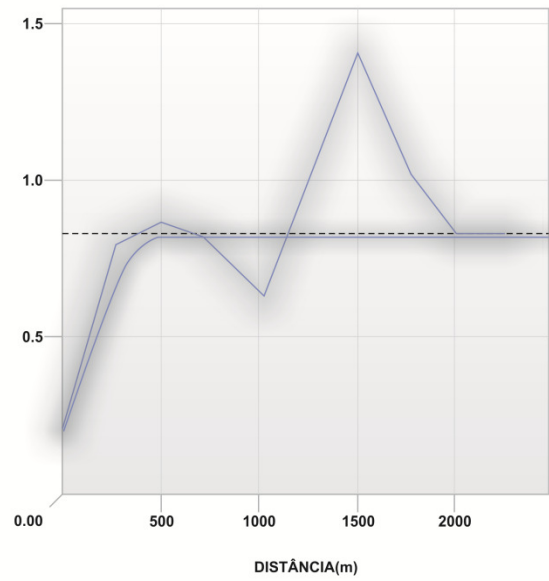
Anexo I

Variogramas Direcionais Ajustados dos Dados Originais

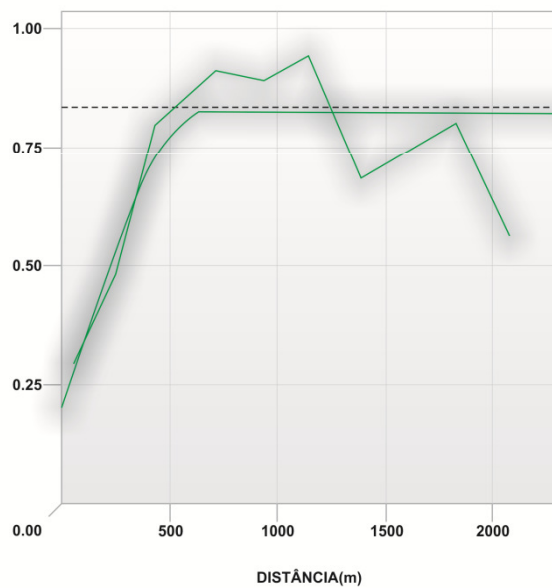
Variogramas Direcionais Ajustados dos Dados Originais Para as direções de 0° (I.1), 45° (I.2), 90° (I.3) e 135° (I.4)



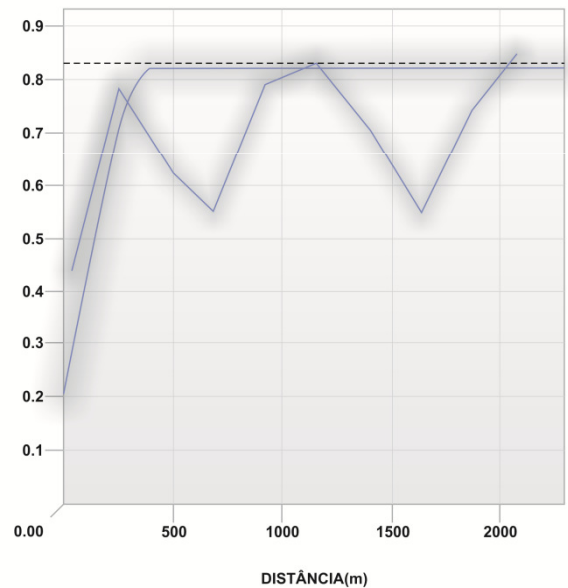
I.1



I.2



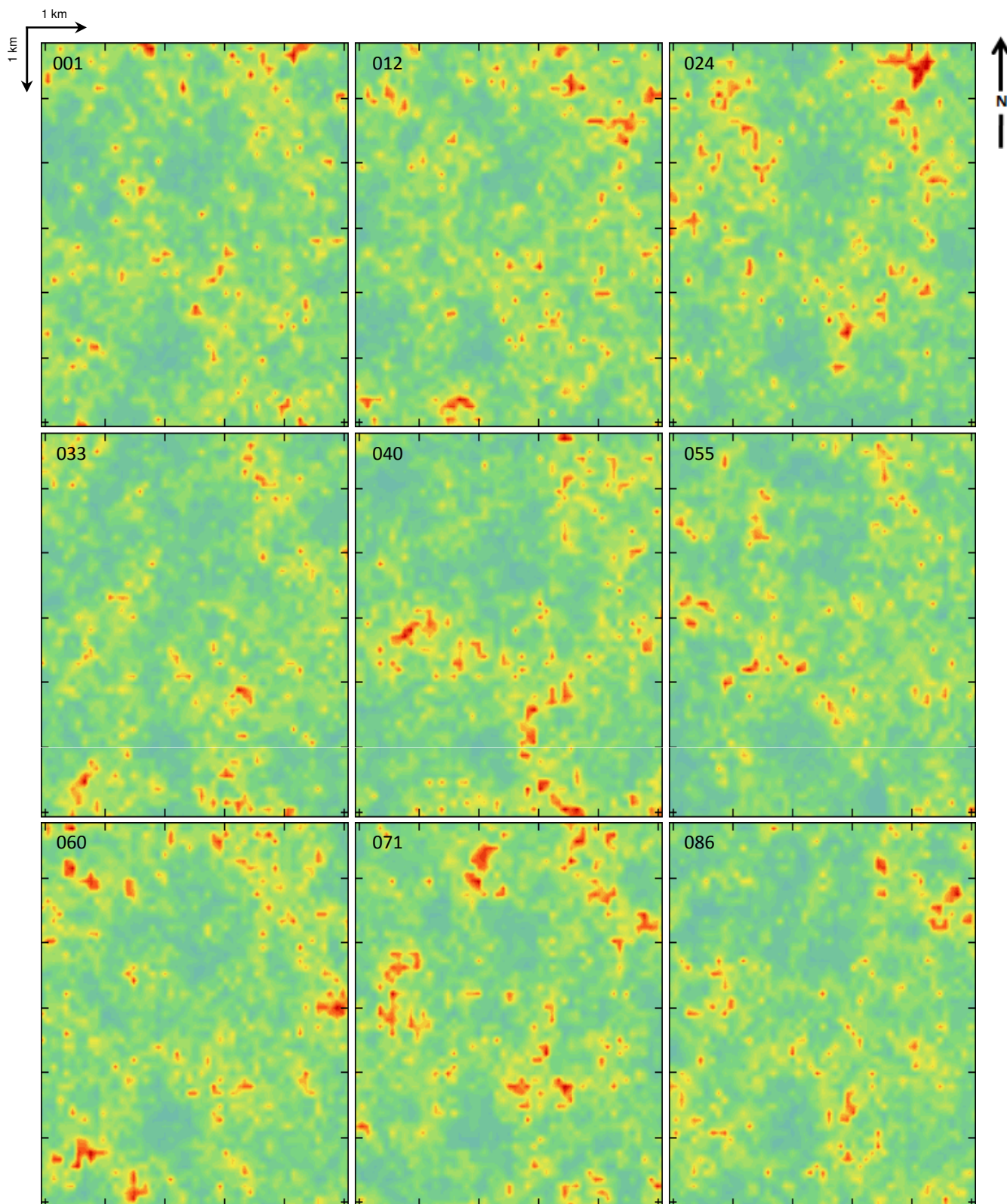
I.3



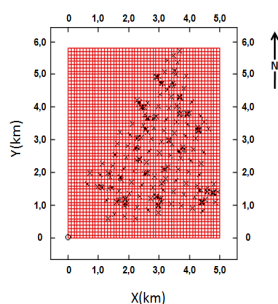
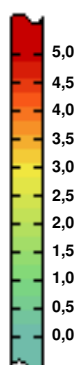
I.4

Anexo II

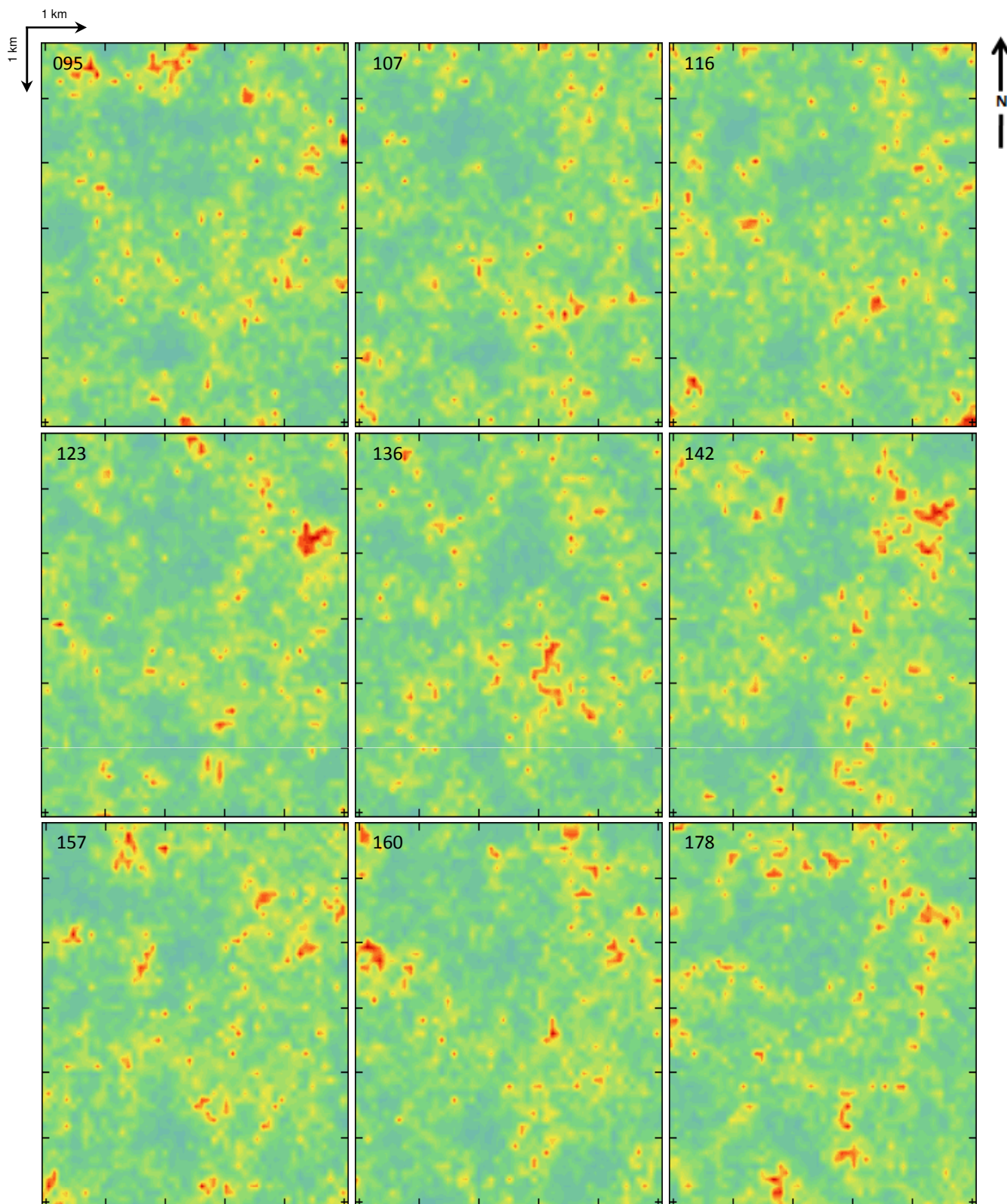
Realizações da ssG



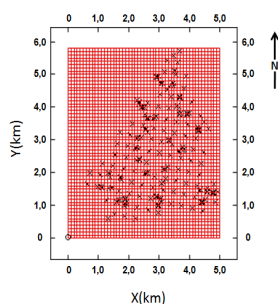
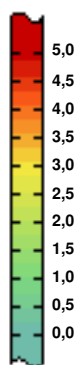
Concentração de Cádmio (ppm)



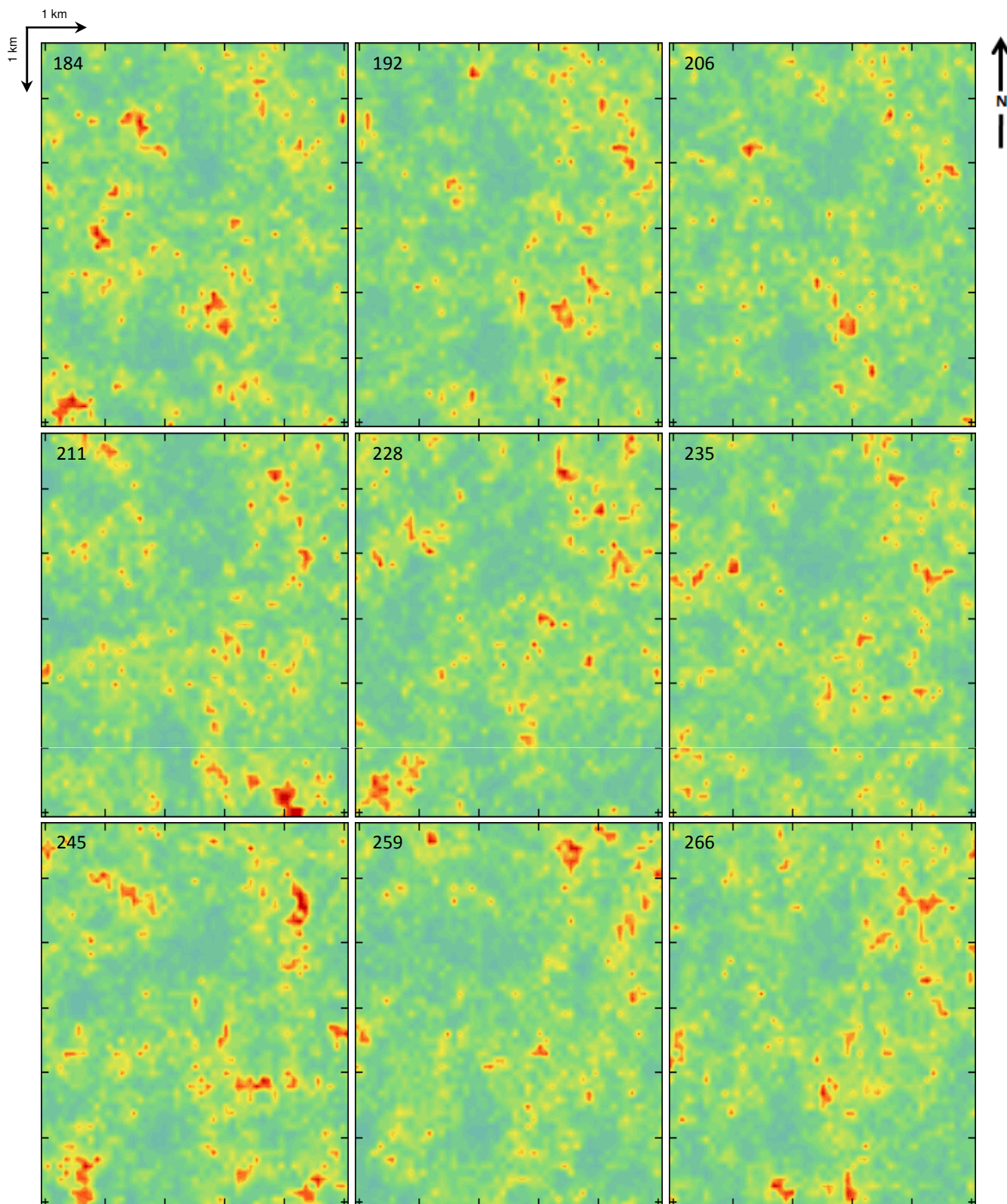
Resultados visuais das simulações estocásticas geradas pelo algoritmo da simulação sequencial Gaussiana, correspondentes às realizações de número 001, 012, 024, 033, 040, 055, 060, 071 e 086 (indicado no canto superior esquerdo de cada imagem).



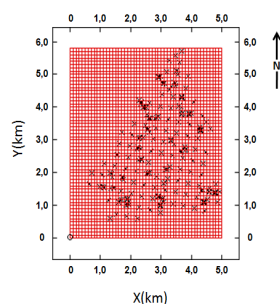
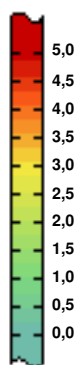
Concentração de Cádmio (ppm)



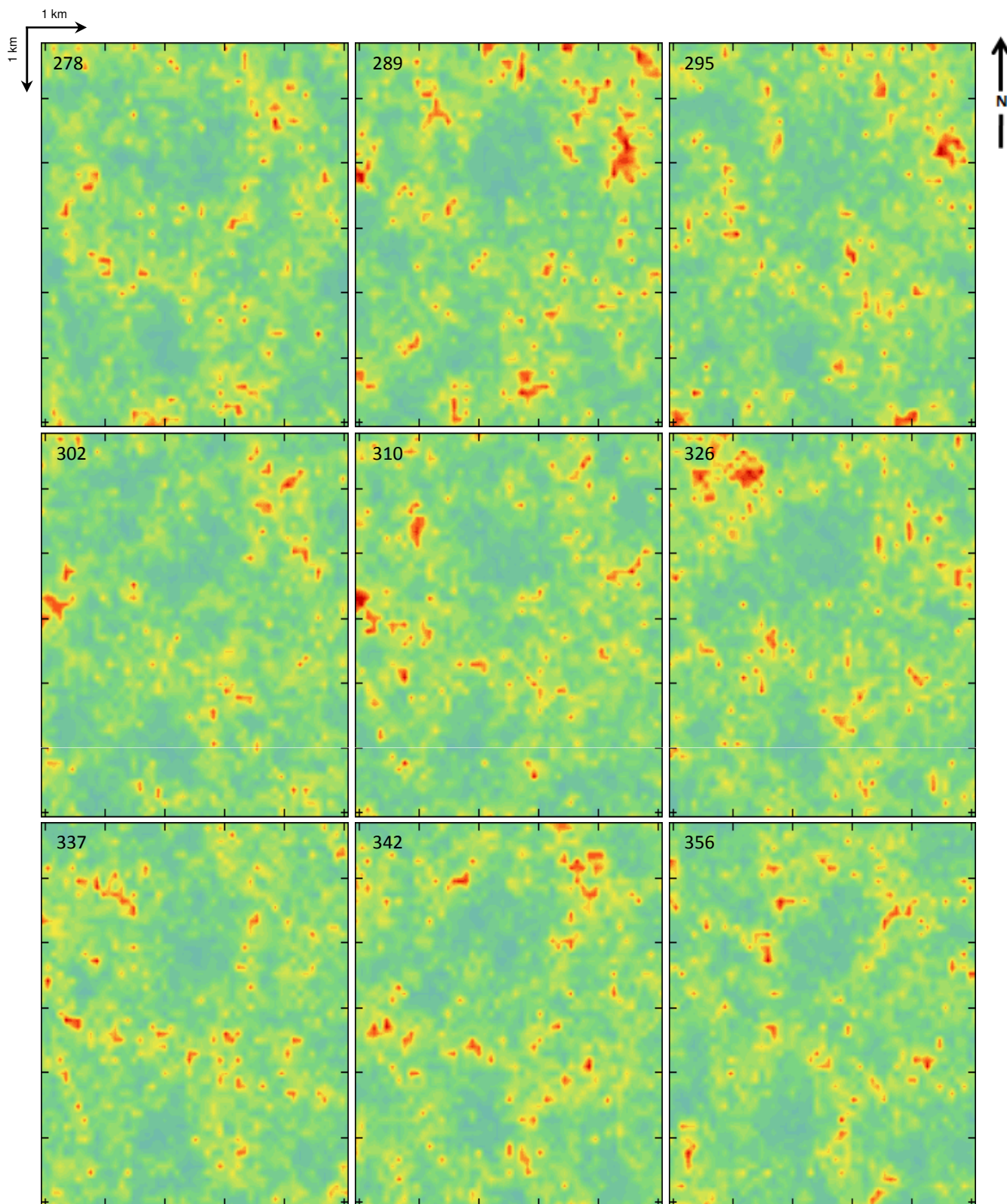
Resultados visuais das simulações estocásticas geradas pelo algoritmo da simulação sequencial Gaussiana, correspondentes às realizações de número 095, 107, 116, 123, 136, 142, 157, 160 e 178 (indicado no canto superior esquerdo de cada imagem).



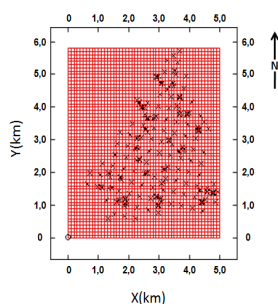
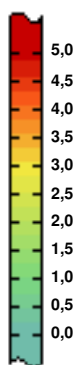
Concentração de Cádmio (ppm)



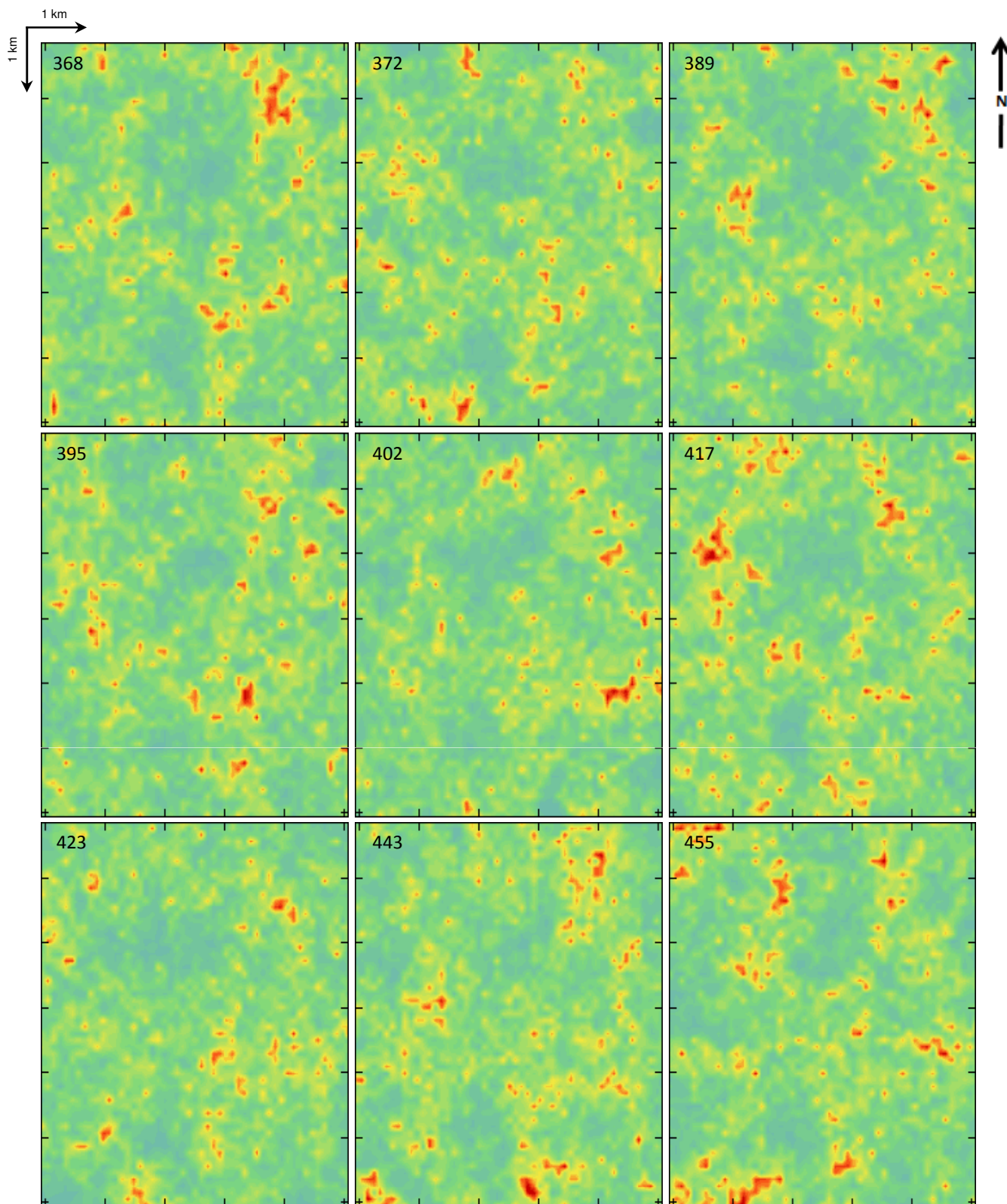
Resultados visuais das simulações estocásticas geradas pelo algoritmo da simulação sequencial Gaussiana, correspondentes às realizações de número 184, 192, 206, 211, 228, 235, 245, 259 e 266 (indicado no canto superior esquerdo de cada imagem).



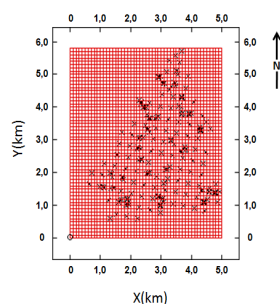
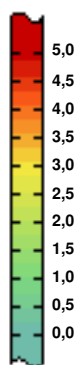
Concentração de Cádmio (ppm)



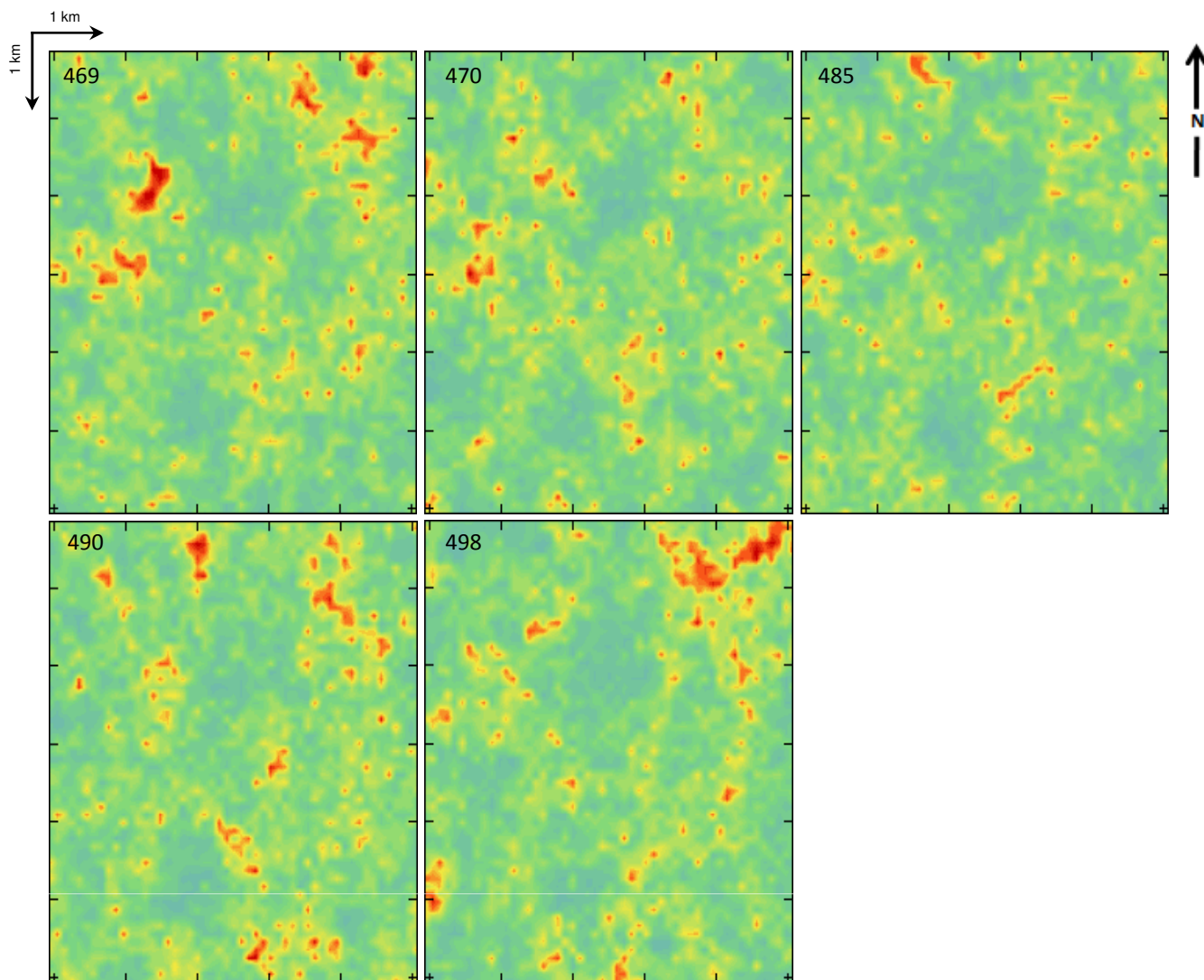
Resultados visuais das simulações estocásticas geradas pelo algoritmo da simulação sequencial Gaussiana, correspondentes às realizações de número 278, 289, 295, 302, 310, 326, 337, 342 e 356 (indicado no canto superior esquerdo de cada imagem).



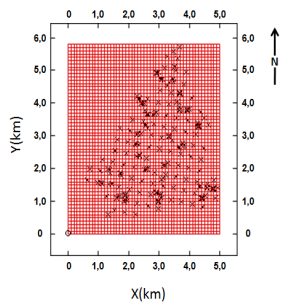
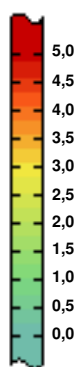
Concentração de Cádmio (ppm)



Resultados visuais das simulações estocásticas geradas pelo algoritmo da simulação sequencial Gaussiana, correspondentes às realizações de número 368, 372, 389, 395, 402, 417, 423, 443 e 445 (indicado no canto superior esquerdo de cada imagem).



Concentração de Cádmio (ppm)



Resultados visuais das simulações estocásticas geradas pelo algoritmo da simulação sequencial Gaussiana, correspondentes às realizações de número 469, 470, 485, 490 e 498 (indicado no canto superior esquerdo de cada imagem).