

## O Uso da Geoestatística na Avaliação dos Parâmetros Hidrogeológicos para Compor o Mapa de Vulnerabilidade Intrínseca de Aquíferos

### *The Use of Geostatistics in the Assessment of Hydrogeological Parameters to Map the Composition of the Intrinsic Vulnerability of Aquifers.*

Érico Gaspar Lisbôa<sup>1</sup>, Jorge Manuel C. Machado de Carvalho<sup>2</sup> e Ronaldo Lopes Rodrigues Mendes<sup>3</sup>

<sup>1</sup> Departamento de Engenharia Civil, Faculdade de Engenharia da Universidade do Porto, Porto, Portugal

dec12006@fe.up.pt

<sup>2</sup> Departamento de Engenharia de Minas / CERENA, Faculdade de Engenharia da Universidade do Porto, Porto, Portugal

jorcarv@fe.up.pt

<sup>3</sup> Núcleo de Meio Ambiente, Universidade Federal do Pará, Belém, PA, Brasil

rlopes@ufpa.br

Recebido: 15/07/15 - Revisado: 16/11/15 - Aceito: 04/12/15

## RESUMO

O objetivo deste trabalho é realizar uma estimativa geoestatística de parâmetros hidrogeológicos para compor o mapa de vulnerabilidade intrínseca de aquíferos. Para tanto, realiza-se uma análise exploratória de dados de poços extraídos do Sistema de Informação de Águas Subterrâneas, os quais se referem aos parâmetros confinamento, litologia e nível estático. Assim, inicialmente, entre outros procedimentos, verificam-se dados inconsistentes, discriminam-se estatísticas e aplica-se o teste de normalidade. Em seguida, faz-se a descrição espacial preliminar dos respectivos parâmetros, em que se identificam padrões, aleatoriedade e agrupamento dos dados de poços, além de tendência para determinadas direções. Após as etapas preliminares, a metodologia geoestatística complementa-se por três fases. A primeira refere-se ao estudo variográfico, caracterizada pela construção de variogramas experimentais. Na segunda fase, faz-se ajuste automático a esses variogramas por três modelos matemáticos: esférico, exponencial e gaussiano. Na sequência, por validação cruzada, estima-se o desempenho dos modelos variográficos ajustados por meio de indicadores estatísticos. Após validação dos modelos variográficos, realiza-se a última fase, em que se faz a estimativa geoestatística por krigagem. Os valores estimados de parâmetros hidrogeológicos, quando multiplicados entre si, compuseram a estimativa do índice de vulnerabilidade intrínseca ilustrado na forma de mapa. Essa proposta foi aplicada numa área que abrange, parcialmente, a cidade de Belém. Os resultados mostraram que a abordagem metodológica consiste em uma excelente ferramenta de planejamento e gestão para proteção das águas subterrâneas.

**Palavras Chave:** Vulnerabilidade. Geoestatística. Águas subterrâneas. Cidade de Belém/PA.

## ABSTRACT

The aim of this study is to perform a geostatistical estimation of hydrogeological parameters to compose the intrinsic vulnerability map for an aquifer. The study presents an exploratory analysis of well data extracted from the Groundwater Information System (SIAGAS), which refer to parameters such as degree of confining, lithology and static level. Therefore, initially, among other procedures, we look for inconsistent data, we produce descriptive statistics and we also apply a normality test to the dataset. The following steps were to make up the primary space description of the respective parameters based on the corresponding identified patterns, randomness and grouping of well data in addition to exploring the existence of a tendency along certain directions. After the preliminary steps, the geostatistical methodology is then applied divided into three phases. The first refers to the variogram study, characterized by the construction of the experimental variograms. In the second phase, automatic adjustment is made to these variograms for three mathematical models: spherical, exponential and gaussian. Then, by cross validation, the performance is estimated for the different models adjusted using statistical indicators. After validation of these models, the last phase takes place, in which we perform the geostatistical estimation by kriging. The estimated values of hydrogeological parameters, when jointly multiplied, composes the estimation of the intrinsic vulnerability index illustrated in the form of a map. This proposal was applied in an area partially covering the city of Belém. The results showed that the methodological approach represents an excellent planning and management tool for the protection of groundwater aquifers.

**Keywords:** Vulnerability. Geostatistical. Groundwater. City of Belém/PA

## INTRODUÇÃO

Segundo o Conselho Nacional de Pesquisa dos Estados Unidos (National Research Council - NRC) (1993), a avaliação da vulnerabilidade intrínseca, apenas, abrange características físicas e climáticas do meio hidrogeológico. Essa definição enquadra-se numa base de conhecimento hidrogeológico geral do transporte de contaminantes em aquíferos, simplificada por métodos que funcionam como sistemas de indexações. Assim, esses sistemas são lineares, que categorizam classes relativas em atributos qualitativos, pela ponderação ou não de parâmetros hidrogeológicos, cujo resultado apresenta-se, usualmente, na forma de índice (e.g. ALLER et al., 1987; FOSTER; HIRATA, 1988; VAN STEMPVOORT et al., 1993; CIVITA, 1994).

A vantagem no uso de índices de vulnerabilidade intrínseca refere-se ao mapeamento dessas informações num domínio espacial. Essas geoinformações são uma das características mais importantes para comunicação ao usuário e, sobretudo, ao tomador de decisão, como ferramenta de gestão, com vistas à proteção dos recursos hídricos subterrâneos. Contudo, geralmente, essas informações não se encontram regularmente distribuídas no domínio espacial.

Utilizar modelos de interpolação apresenta-se como uma técnica para contornar essa limitação, uma vez que se consegue estimar valores dessas informações numa área não amostrada, isto é, onde tais informações não estão disponibilizadas. Para efetuar esse procedimento, de natureza matemática, ajusta-se uma função aos pontos não amostrados com base em valores obtidos em pontos amostrados. Dentre os modelos de interpolação, têm-se os não geoestatísticos (e.g. inverso do quadrado da distância, média móvel, entre outros); e, os de natureza geoestatística (e.g. krigagem ordinária). A diferença mais marcante entre os diferentes tipos de interpoladores resume-se na forma de ponderar ou de atribuir pesos ao conjunto de observações no sentido de estimar o valor da variável de interesse em pontos diversos dos efetivamente amostrados. Complementarmente, a possibilidade de estimar, por construção metodológica, o campo de incertezas (ou de variâncias) associado à estimativa da variável de estudo, distingue a abordagem geoestatística em relação a outros procedimentos de natureza distinta. A krigagem ordinária simples, por exemplo, é similar à interpolação por média móvel ponderada, exceto que os pesos são determinados por uma análise espacial por um variograma experimental, fornecendo estimativas não tendenciosas e com variância mínima.

Assim, a aplicação de propostas geoestatísticas foi desenvolvida para mapear a vulnerabilidade (ASSAF; SAADEH, 2009; RUSSO, 2009; BAALOUSHA, 2010; LISBÔA et al., 2013; FERREIRA, 2014). No entanto, nessas propostas, é comum admitir que a distribuição espacial da vulnerabilidade face aos parâmetros hidrogeológicos, seja realizada com o suporte de análise geoestatística. Essa consideração pode vir a alterar a estimativa da distribuição da vulnerabilidade no domínio espacial. Logo, a proposta deste artigo é realizar estimativa geoestatística de parâmetros hidrogeológicos para compor o mapa de vulnerabilidade intrínseca.

Para tanto, escolheu-se o método de avaliação GOD (FOSTER; HIRATA, 1988), opção justificada pelo fato de que

dados hidrogeológicos que são requeridos por outros sistemas, e.g. método DRASTIC (ALLER et al., 1987), não se encontraram diretamente disponibilizados (e.g. topografia, condutividade hidráulica). O índice GOD considera a inacessibilidade hidráulica e a capacidade de atenuação dos estratos de cobertura com base na aplicação de três parâmetros hidrogeológicos (grau de confinamento, estratos de cobertura e profundidade do nível de água). O método DRASTIC, por exemplo, demandaria sete planos de informações, incluindo profundidade do lençol freático, recarga, condutividade hidráulica, mapa de pedocobertura, que não se encontram disponibilizada em sua integralidade no âmbito da base de dados do SIAGAS (Sistema de Informação Geográfica de Águas Subterrâneas).

Logo, na análise, empregam-se estatísticas, faz-se o teste de normalidade e verificam-se dados inconsistentes. Posteriormente, faz-se a descrição espacial preliminar dos parâmetros hidrogeológicos. Após as etapas preliminares, a metodologia geoestatística complementa-se por três fases: estudo variográfico, ajuste automático e desempenho da modelagem variográfica. O processo de validação dos modelos variográficos integra a última fase metodológica, onde se faz a estimativa geoestatística por krigagem. Essa proposta foi aplicada numa área que abrange, parcialmente, a cidade de Belém, cuja caracterização se descreve a seguir.

## CARACTERIZAÇÃO DA ÁREA DE ESTUDO

O domínio espacial foi delimitado em 1.64 km<sup>2</sup>, onde se situam 18 bacias hidrográficas que compõem a cidade de Belém e ilha de Outeiro, que integram a região metropolitana de Belém (RMB) (Figura 1).

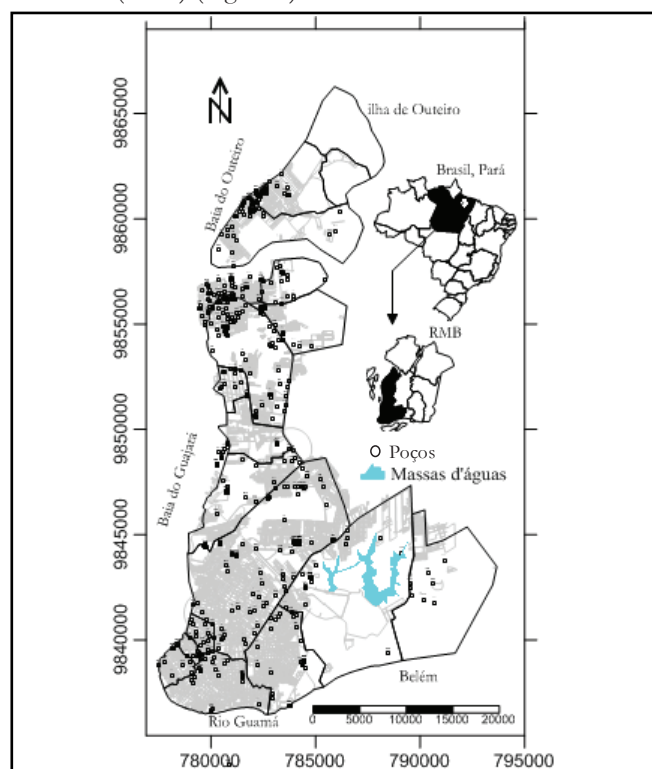


Figura 1 – Domínio espacial da área estudada

Quadro 1 – Litoestratigrafia e hidrogeologia de Belém

E	P	EP	U	e	SL	C.H
Cenozóica	Quaternário	Holoceno-Pleistoceno	Clásticos-Friáveis	10 m	Constituem-se de areias variadas e argilas, distribuída ao longo dos rios, praias e zonas de mangues.	Potencialidade hidrogeológica alta para bateria de poços rasos, boas condições de recarga. Vazão de 10m <sup>3</sup> /h.
						Terciário
	Barreiras	70 m	Constituído por argilas vermelhas com níveis caulinizados e areias continentais	Média potencialidade hidrogeológica. Satisfaz demandas de 20 a 80m <sup>3</sup> /h. Alto teor de ferro.		
	Oligoceno	Pirabas	Superior	80 m	Constituem a parte marinha da bacia, é caracterizado por calcário, arenitos calcíferos, folhetos e abundante conteúdo fóssilífero.	
			Inferior	120m	Caracterizado por arenitos de granulometria fina a grossa e pelitos associados. São frequentes níveis conglomeráticos	Excelente potencial hidrogeológico. Vazões superiores a 200m <sup>3</sup> /h.
-	-	-	Pré-pirabas	> 300m	Arenitos finos com intercalações de siltos e argila.	Hidrodinâmica desconhecida. Camadas arenosas de 20 m de espessura.

Em que E – Era; P – Período; EP – Época; U – Unidades; e – Espessura; S.L – Síntese Litológica; C.H – Características hidrogeológicas

Essa região, pertencente ao estado Pará, norte do Brasil; limitada ao sul e a oeste, pelo rio Guamá e baía de Guajará. Os dados de 539 poços foram obtidos pela rede de monitoramento do SIAGAS. Com uma densidade de  $0,4 \times 10^{-3}$  poços/km<sup>2</sup>, esses dados foram interpretados e utilizados para determinação das características litológicas do aquífero, e distância do nível d'água até a superfície. Nesses termos, o Quadro 1 sintetiza as informações sobre as formações geológicas, e litológicas que podem constituir aquíferos inseridos no arcabouço geológico da área.

A geologia de grande parte do domínio espacial pertence ao período quaternário pleistoceno, de formação pós-barreiras, cujos sedimentos são constituídos por areias amarelo-alaranjadas de granulometria fina a média. A outra parte, formada por sedimentos holoceno, caracteriza-se por sedimentos inconsolidados constituído de areias, siltes e argilas, distribuídos na faixa costeira, leitos de rios e manguezais. Matta (2002) ainda referiu sobre uma estreita faixa que aflora a formação Barreiras (Figura 2).

O Mapa Hidrogeológico da região está ilustrado pela Figura 3, e baseou-se nos estudos realizados pela Companhia

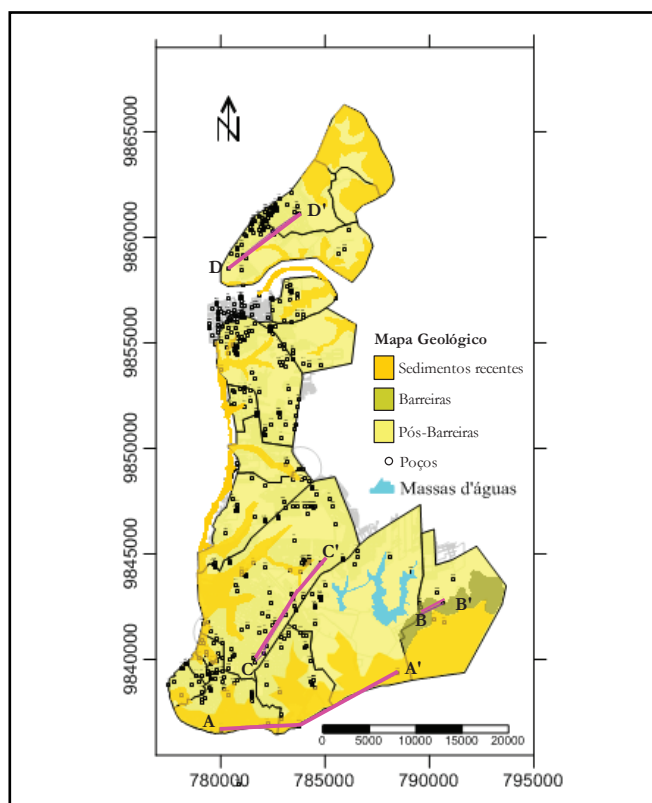


Figura 2 – Mapa Geológico do domínio espacial

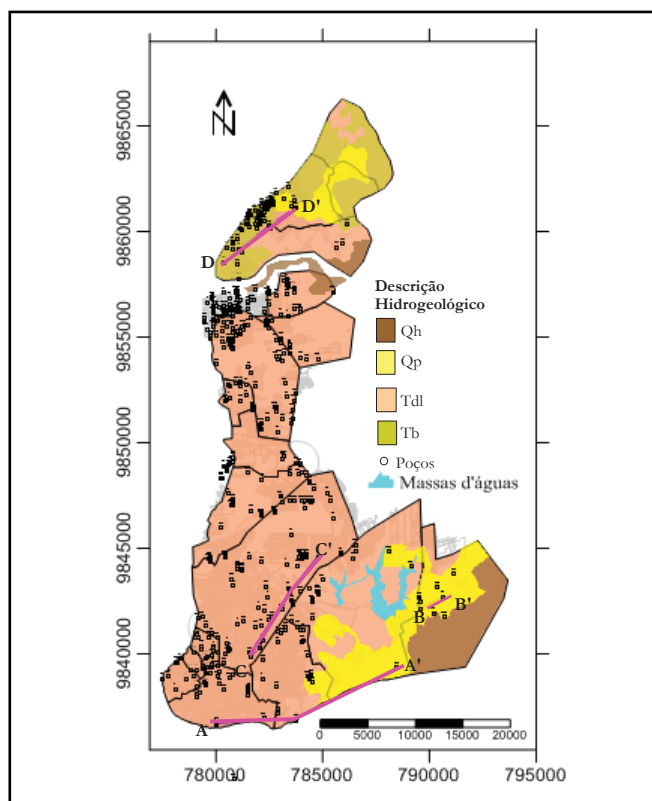


Figura 3 – Mapa Hidrogeológico do domínio espacial

de Pesquisa e Recursos Minerais (CPRM, 2001). Esses estudos referiram que, por se tratar de uma região sedimentar, cuja

capacidade de armazenamento e circulação de água dependem da porosidade das rochas das unidades Pirabas, Barreiras e das coberturas detriticas lateriticas e aluvionares, a compartimentação hidrolitológica é composta de aquíferos: intergranulares descontínuos, livres; intergranulares contínuos, livres e semiconfinados; e, intergranulares contínuos, livres e confinados.

Em que a sigla Qh representa aquíferos porosos descontínuos de extensões variáveis e livres, e de permeabilidade média a alta. O aproveitamento é feito por poços de até 18 metros.

Fornecem vazões da ordem de 10m<sup>3</sup>/h, são muito produtivos e de grande importância hidrogeológica. O termo Qp representa aquíferos porosos descontínuos de extensões variáveis e livres, contudo, possuem alta permeabilidade. Seu aproveitamento é feito através de poços de até 12 metros que permitem vazões bastante elevadas, na ordem de até 10m<sup>3</sup>/h.

Por outro lado, o termo Tdl refere-se a aquíferos porosos descontínuos, extensões variáveis, semiconfinados a livre; de permeabilidade baixa a média. O aproveitamento é obtido por poços de até 40 metros. Possuem grande importância hidrogeológica e oferecem vazões de 1 a 3m<sup>3</sup>/h. Por sua vez, Tb refere-se a aquíferos porosos semiconfinados a confinados e extensão regional, de permeabilidade média a alta. O aproveitamento é obtido por poços de até 100m de profundidade. Possuem grande importância hidrogeológica e apresentam vazões na ordem de até 100 m<sup>3</sup>/h. Detalhes sobre a hidrogeologia da região pode ser consultado em Matta (2002); Palheta (2008).

Com base no mapa geológico e hidrogeológico, verifica-se que mais de 80% das amostras de poços estão inseridos no aquífero da formação Pós-Barreiras. Entretanto, pouco mais de 74% dos poços captam água de aquífero livre; e, o restante de aquíferos semiconfinados a confinados. O estudo de CPRM (2001) considerou que os aquíferos confinados do domínio espacial em questão apresentam-se naturalmente protegidos, como os aquíferos da formação Barreiras e Pirabas. Assim sendo, analisaram-se apenas aquíferos livres, especialmente os das formações Quaternários e Barreiras, i.e. parâmetro  $x_G = 1$ .

No entanto, para quantificar o parâmetro relativo à ocorrência litológica, considerou-se o perfil geológico e litoconstrutivo de poços. Assim, ilustram-se quatro seções A-A', B-B', C-C' e D-D'.

A seção A-A' é composto de quatro poços, perpassa pela formação geológica de sedimentos recentes, e abrange aquíferos da formação Pirabas e Pós-Barreiras. A seção B-B' perpassa pela formação geológica Barreiras, e abrange aquífero da formação Pós-Barreiras. A seção C-C', por sua vez, perpassa pela formação geológica Pós-Barreira, e abrange aquíferos da formação Pirabas. A seção D-D' perpassa pela formação geológica Barreiras, e abrange aquíferos da formação Pós-Barreiras (Figura 4).

Portanto, com base no perfil litoconstrutivo dos 539 poços foi possível classificar os materiais que constituem a zona vadosa.

## METODOLOGIA

Para alcançar o objetivo do trabalho, recorre-se a catalogação de dados de poços, advindos da base de dados SIAGAS. Esses dados foram utilizados para interpretar os parâmetros que compõem o método GOD: ocorrência litológica ( $x_o$ ) e a distância até o lençol freático ( $x_D$ ). Assim,  $x_o$  e  $x_D$ , e ainda o parâmetro  $x_G = 1$ , quando multiplicados entre si, compõem as considerações de vulnerabilidade, proposto por Foster e Hirata (1988).

Assim, realiza-se a análise exploratória de uma amostra de 539 dados de poços, para o domínio espacial considerado. Faz-se uma análise da estatística descritiva referente aos respec-

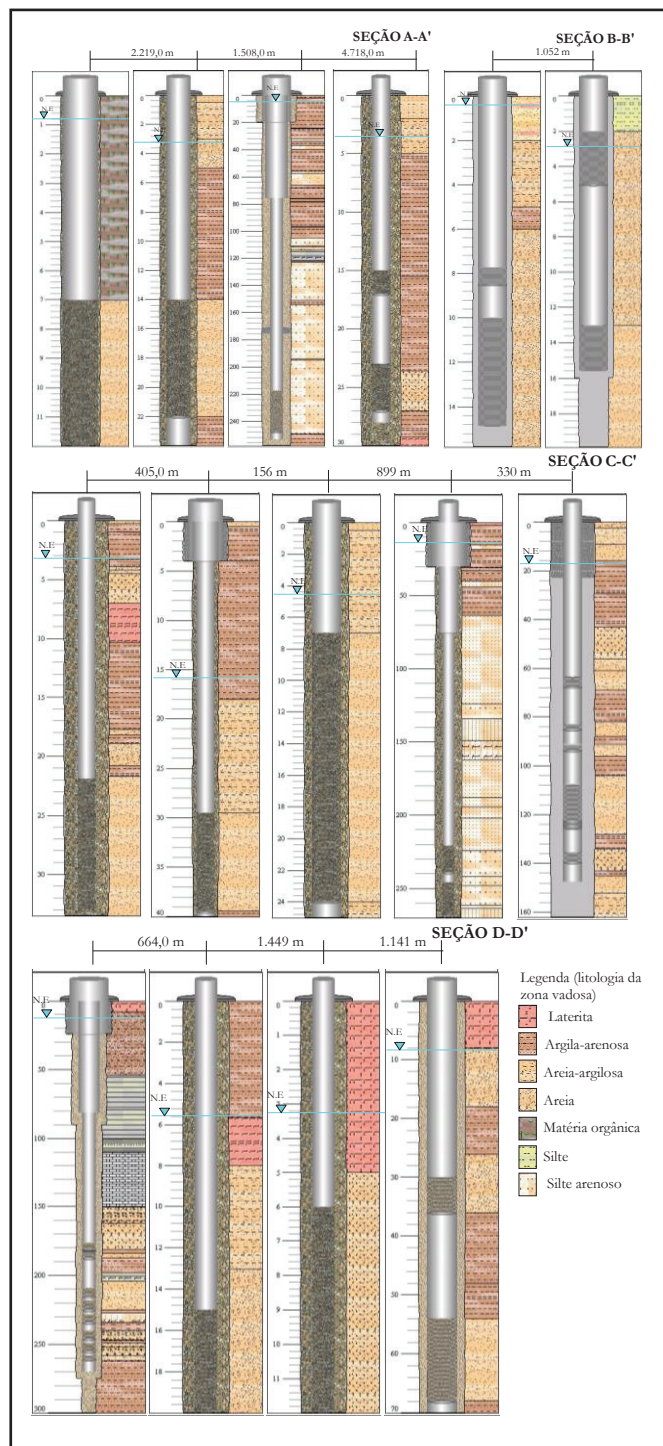


Figura 4 – Perfil litoconstrutivo de poços: Seção A-A'; B-B'; C-C'; e, D-D'

tivos parâmetros hidrogeológicos; e, verifica-se a normalidade por teste do tipo Komolgorov-Smirnov. Esse teste é precedido da verificação da presença de outliers pelo gráfico Box-plot. Caso detectado, excluem-se outliers pois podem interferir nos testes estatísticos.

Em seguida, caso  $x_o$  e  $x_D$  não apresentem-se normalmente distribuídos; aplica-se a transformação de Box-Cox. Para esse procedimento utiliza-se o suplemento Action<sup>®</sup>, instalado ao MS-Excel<sup>®</sup>; e, em seguida, faz-se, novamente, o teste de normalidade.

Na sequência realiza-se a análise da descrição espacial preliminar via método Minimum curvature, pelo *software* Surfer 9<sup>®</sup>; para detectar padrões sistemáticos ou aleatórios da distribuição dos dados  $x_o$  e  $x_D$  no domínio espacial considerado. Justifica-se a escolha desse método, pois o mesmo proporciona uma exibição da aleatoriedade, agrupamento ou regularidade de distribuição.

Efetua-se análise de deriva pela ferramenta polytool do *software* MatLab<sup>®</sup>. Assim, permite-se representar  $x_o$  e  $x_D$  segundo os eixos cartesianos, efetuando uma regressão linear. Desse modo, verifica-se, imediatamente, se os dados variam, em média, num comportamento bem definido, segundo uma determinada direção. Caso as retas dessa regressão apresentem uma inclinação acentuada, assume-se que os parâmetros em estudo apresentam o fenômeno de deriva, contrariando as hipóteses de estacionaridade.

Após a análise exploratória de dados, bem como a descrição espacial preliminar; a metodologia geoestatística se complementa pelo cumprimento de três fases: (i) estudo variográfico; (ii) ajustes ao variograma experimental; e, (iii) estimação geoestatística.

Na primeira fase, analisa-se, de imediato; o mapa variográfico de pixel produzido via *software* Variowin<sup>®</sup>. Esse procedimento é adotado para detectar a ocorrência de anisotropias, bem como determinar o eixo e ângulo de maior variabilidade espacial. Com base nessas direções, constroem-se os variogramas direcionais e omnidirecionais via Variowin<sup>®</sup>. O propósito da construção do variograma omnidirecionais é ratificar a existência de tendência (deriva), anteriormente analisada pelas retas de regressão.

Sobrepoem-se ambos os variograma para detectar ocorrências ou não de continuidade espacial desigual em uma ou mais direções. A elaboração desses variogramas e mapas variográficos foram precedidos das considerações acerca da distância entre os pares de pontos (*lag*), número e tolerância de *lag*. Portanto, com base no domínio espacial, na distribuição de dados e análise preliminar de tendência da continuidade espacial; define-se o produto da distância entre os pares de pontos e o número de *lags*.

Como os dados estão irregularmente distribuídos, analisam-se os variogramas, considerando uma tolerância angular de 45°; adotando-se um valor para maximum bandwidth. Assim sendo, inicia-se a segunda fase, i.e. realiza-se ajustes matemáticos aos variogramas. Para tanto, adota-se modelos SPH, EXP e GAU, definidos, respectivamente:

$$\gamma(h) = C_0 + C_1 \left[ 1,5 \left( \frac{h}{a} \right) - 0,5 \left( \frac{h}{a} \right)^3 \right] \quad (1)$$

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

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

em que  $C_0$  é o efeito pepita,  $C_1$  é a diferença entre o patamar  $-C_0$  e  $C_0$ ; e,  $a$  é o alcance. Em seguida, realizam-se ajustes automáticos aos variogramas. A qualidade dos ajustes foi calculada pelo índice IGF (Indicative Goodness of Fit), proposto por Pannatier (1996), expresso por:

$$IGF = \frac{1}{N} \sum_{k=1}^N \sum_{i=0}^{n(k)} \frac{P(i)}{\sum_{i=0}^{n(k)} P(i)} \cdot \frac{D(k)}{d(i)} \cdot \left[ \frac{\gamma(i) - \gamma'(i)}{\sigma^2} \right] \quad (4)$$

em que  $N$  é o número de variogramas direcionais;  $n(k)$  igual ao número *lags* relativo ao variograma  $k$ ;  $D(k)$  é a distância máxima;  $P(i)$  igual número de pares para o passo  $i$  do variograma  $k$ ;  $d(i)$  é igual a distância média dos pares para o passo  $i$  do variograma  $k$ ;  $\gamma(i)$  é medida experimental da continuidade espacial para o passo  $i$ ;  $\gamma'(i)$  é medida modelada da continuidade espacial para  $d(i)$ ; e,  $\sigma^2$  é igual a ( $C_1$ ) variância dos dados para o variograma (cruzado).

Considerando os modelos mais bem ajustados, operam-se as possíveis correções anisotrópicas pelo *software* Variowin<sup>®</sup>. Para correção da anisotropia geométrica, o ângulo  $\theta$  entre o norte e o eixo maior da elipse, que representa a maior continuidade espacial; é rotacionado. Conforme descrito por Yamamoto (2001), após essa rotação, dado um vetor  $b = (bx; by)$  entre dois pontos quaisquer; obtem-se um novo vetor  $b' = (b'x; b'y)$  por:

$$\begin{bmatrix} h'x \\ h'y \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \cdot \begin{bmatrix} hx \\ hy \end{bmatrix} \quad (5)$$

Após a rotação, o redimensionamento da elipse é representado pelo raio de menor continuidade espacial. Assim, feito a correção, o variograma da direção de menor continuidade é utilizado como o variograma anisotrópico. Caso seja detectado anisotropia zonal, faz-se a correção por meio de um variograma direcional, equivalente com a distância reduzida, considerando  $C$ , o maior valor de patamar apresentado entre os variogramas direcionais construídos de tal modo que:  $\gamma'(b) = C \cdot \gamma(b/a)$ . Onde  $C$  e  $a$ , são os valores de patamar e alcance do variograma na direção que apresentou a anisotropia zonal.

A anisotropia mista ou combinada pode ser corrigida por:  $\gamma'(b) = C_1 \cdot \gamma_1(b') + C_2 \cdot \gamma_2(b')$ . Em que  $C1$  representa o patamar do variograma direcional de maior alcance; e,  $C2$  o que apresenta menor alcance. O valor de  $b'$  é igual a:

$$h' = \left[ \left( \frac{x}{a_x} \right)^2 + \left( \frac{y}{a_y} \right)^2 \right] \quad (6)$$

em que  $a_x$  e  $a_y$  representam os alcances nas direções  $x$  e  $y$ , respectivamente. Detalhes desse procedimento podem ser consultados em Soares (2000).

Corrigidas anisotropias fazem-se estimativas geoestatísticas através da krigagem ordinária pelo modelo corrigido.

Caso se detecte o fenômeno de deriva, utiliza-se krigagem de tendência. Assim sendo, avalia-se e compara-se o desempenho dos modelos variográficos ajustados e corrigidos. Em seguida, utiliza-se a técnica de validação cruzada, permitindo uma comparação entre valores estimados e reais, validando o modelo variográfico. Para tanto, faz-se uma análise estatística da diferença entre valores estimados e reais, nomeadamente os resíduos. Calculam-se, então, o erro médio (MSE) e a raiz quadrada do erro quadrático médio (RMSE) por:

$$MSE = \frac{1}{N(h)} \cdot \sum_{i=1}^{N(h)} [Z^*(x_i) - Z(x_i)] \quad (7)$$

$$RMSE = \sqrt{\frac{1}{N(h)} \cdot \sum_{i=1}^{N(h)} [Z^*(x_i) - Z(x_i)]^2} \quad (8)$$

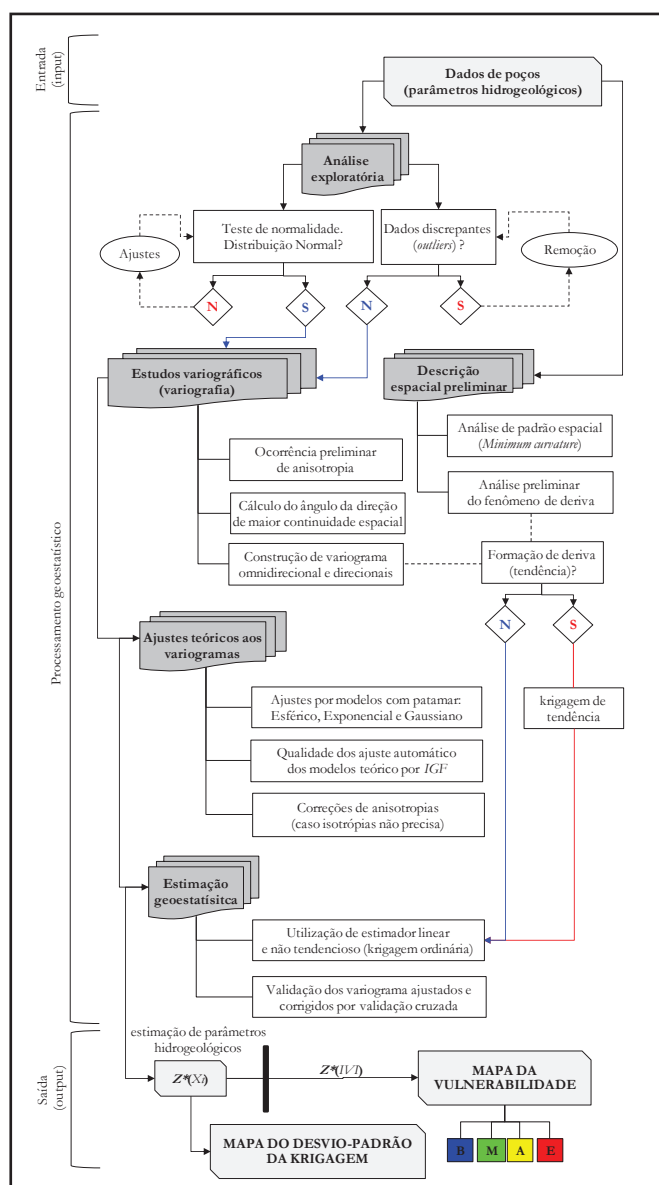


Figura 5 – Esquema da metodologia adotada

em que  $Z^*(x_i)$  é o valor estimado; e,  $Z(x_i)$  é o valor calculado; e refere-se aos parâmetros hidrogeológicos. E,  $N(h)$  é o número de pares ordenados.

Validado o modelo variográfico que apresentou melhor desempenho, faz-se o mapeamento da distribuição espacial de  $x_0$  e  $x_D$  por estimação geoestatística através da krigagem ordinária, acompanhado do mapa do desvio-padrão da Krigagem -  $[Z^*(x_i) - Z(x_i)]^2$  (áreas que há um desvio-padrão muito alto merecem um adensamento da amostragem). E, por fim, compõem-se o mapa de vulnerabilidade:  $Z^*(x_{IV1}) = Z^*(x_0) \times Z^*(x_D)$ . A Figura 5 ilustra o processo metodológico adotado.

## RESULTADOS E DISCUSSÃO

### Análise exploratória de $x_0$ e $x_D$

A análise de dados foi realizada para cada um dos parâmetros hidrogeológicos utilizados na pesquisa. O parâmetro  $x_0$  relacionado à ocorrência litológica é qualitativo, pois se refere a uma classificação da natureza geológica das rochas: não consolidada (sedimentos) ou consolidada (rochas porosas ou duras). Para cada uma dessas classificações, Foster e Hirata (1988) atribuíram valores de 0,4 a 1; requerendo uma identificação litológica. A importância da litologia na avaliação da vulnerabilidade relaciona-se à capacidade de atenuação de contaminantes. Tal capacidade depende da distribuição granulométrica e composição mineralógica na zona não suturada ou camada confinante. Considerando a formação geológica da área investigada como sendo de natureza sedimentar, para a composição litológica argilo-arenosa e areno-argilosa adotou-se o valor de  $x_0$  como sendo igual a 0,55 e 0,65, respectivamente. Concreções ferruginosas e lateritas foram associado ao índice 0,7. Assim, com base na interpretação do perfil litoconstrutivo, obteve-se:

Tabela 1 – Distribuição dos dados de  $x_0$

Ocorrência litológica	Nº poços	%	$x_0$
Argilas	200	37,11	0,4
Argila-arenosa	116	21,52	0,55
Silte	7	1,30	0,60
Areia-argilosa	125	23,19	0,65
Areia	91	16,88	1

A ocorrência litológica foi associada a extratos de cobertura não consolidado da zona vadosa. Por essa consideração, as características litoconstrutivas dos poços analisados revelou que pouca mais de 37% e 21% são argilas e argilas-arenosas; e, areia e areias-argilosas em cerca de 16% e 23%, respectivamente.

O parâmetro  $x_D$  foi relacionado à distância até o lençol freático, sendo de natureza discreta -  $x_D$ (m). Com objetivo de normalizar os resultados do índice, Foster e Hirata (1988) estabeleceram valores de 0,6 a 1, categorizando intervalos de profundidades (Tabela 2).

A maior frequência foi associada a profundidades entre 5 a 20 m, seguido de valores < 5 m. A síntese da análise estatística de  $x_0$  e  $x_D$  está demonstrada na Tabela 3.

Tabela 2 – Distribuição dos dados de  $x_D$

Nível estático (m)	Nº poços	%	$x_D$
Todas as profundidades	25	4,64	1
< 5	159	29,50	0,9
Entre 5 a 20	327	60,67	0,8
Entre 20 a 50	28	5,19	0,7

Tabela 3 – Sumário estatístico de  $x_O$  e  $x_D$  (m)

Estatística descritiva	$x_O$	$x_D$ (m)
Média	0,54	7,58
Desvio-padrão	0,12	5,42
Mínimo	0,40	0,00
Máximo	0,85	34,94
Moda	0,40	6,00
Mediana	0,55	6,00
Variância	0,02	38,53
Coefficiente de variação	0,23	0,71

Apenas os parâmetros  $x_D$  (m), enquadrou-se numa distribuição normal, ratificado pelo teste de Kolmogorov-Smirnov, em que  $t_{K-S}$  calculado igual a 0,017 foi menor que  $t_{K-S}$  tabelado (Figura 6). Embora  $x_O$  e  $x_D$  não tenham apresentado distribuição normal, a transformação via método Box-Cox, não foi realizada por dois motivos: (i) a aplicação desse método fica restrita para valores nulos; e, (ii) qualquer outro método que transforma dados para que a distribuição seja normalmente distribuída; normalizam os resíduos, havendo assim redundância, posto que os valores já apresentam-se normalizados.

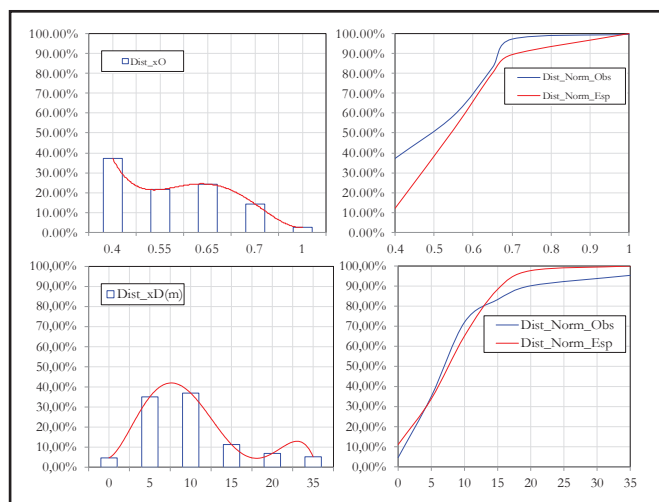


Figura 6 – Distribuição de frequência dos parâmetros  $x_O$

As amostras de  $x_O$  não apresentaram outliers. As amostras de  $x_D$  (m), não apresentaram outliers inferiores; mas, foram detectados 38 outliers superiores (Figura 7).

Para  $x_D$  (m) os dados outliers foram removidos, e novo teste  $t_{K-S}$  foi procedido. O resultado do teste  $t_{K-S}$  permaneceu inalterado, i.e.  $x_D$  (m) enquadrou-se como uma distribuição normal (Figura 8)

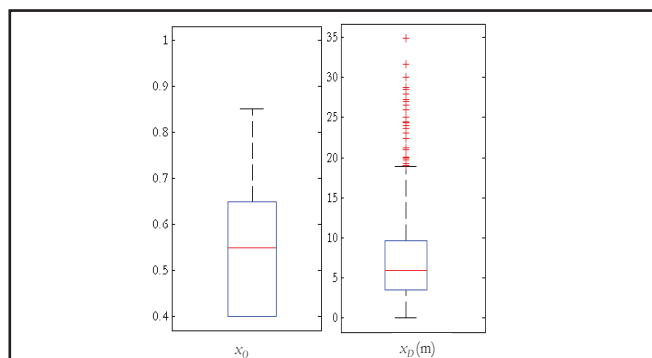


Figura 7 – Box-plot dos parâmetros  $x_O$  e  $x_D$  (m)

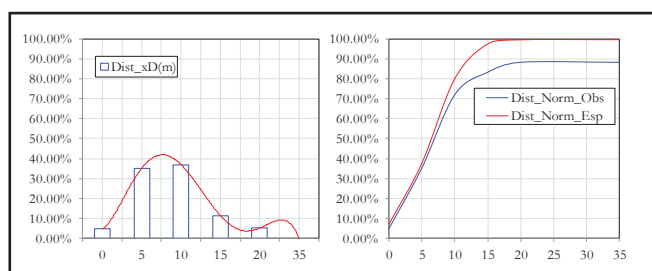


Figura 8 – Distribuição de frequência dos parâmetros

Após a retirada de outliers do conjunto amostral de  $x_D$  (m), o sumário estatístico está descrito pela Tabela 4.

Tabela 4 – Sumário estatístico de  $x_O$  e  $x_D$  (m)

Estatística descritiva	$x_D$ (m)
Média	6,35
Desvio-padrão	4,30
Mínimo	0,00
Máximo	18,80
Moda	6,00
Mediana	6,00
Variância	18,48
Coefficiente de variação	0,68

Portanto, para os estudos geoestatísticos admitiu-se que  $x_O$  e  $x_D$  (m) apresentaram distribuição normal. Pressuposto razoável, posto que o comportamento desses parâmetros aproximou-se dessa condição.

### Descrição espacial preliminar dos parâmetros $x_O$ e $x_D$

A descrição espacial preliminar do parâmetro  $x_O$ , ao analisar os perfis litoconstrutivos dos 539 poços disponíveis, indicou que ao norte da área analisada, o valor entre 0,6 a 0,7 configura-se num padrão argilo-arenoso a arenoso. Por outro lado, ao centro e a sudoeste, não há um padrão definido, de tal modo que o perfil litoconstrutivo apresenta-se ora argilo-arenosa, ora argilosa. Ao Sudeste, esse padrão indicou predomínio de material argiloso (Figura 9A).

A descrição espacial preliminar de  $x_D$  indicou que predominaram, ao norte, noroeste, sul, e sudoeste da área

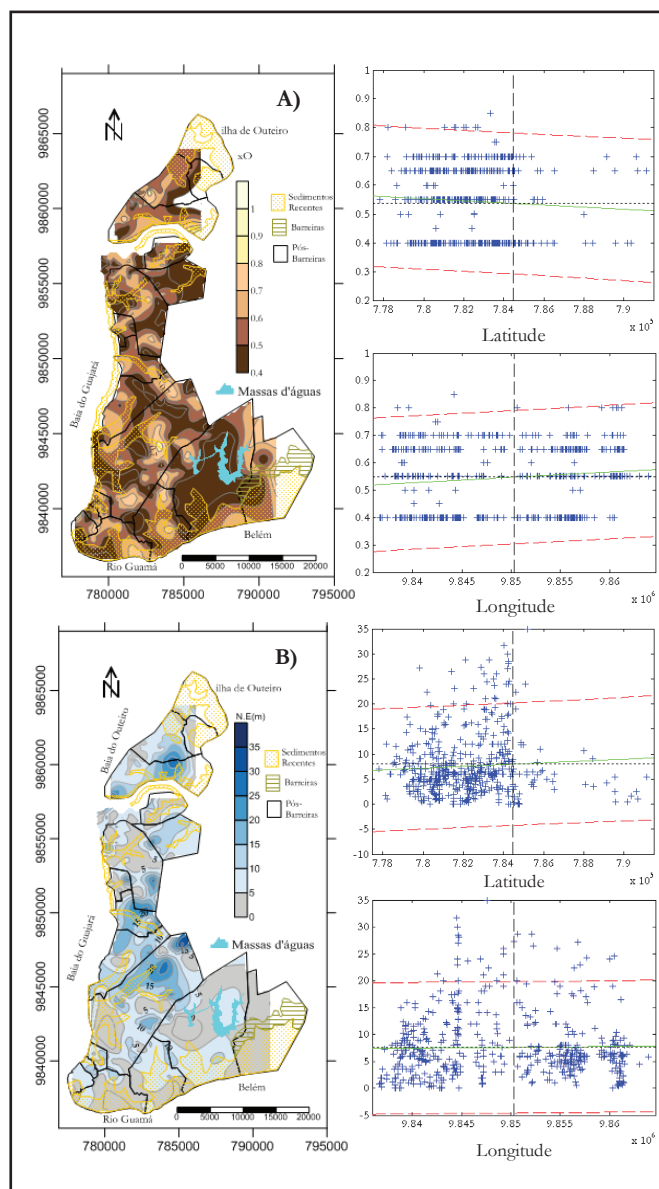


Figura 9 – Interpolação não geoestatística e análise de deriva da distribuição espacial de: A)  $x_0$ ; e, B)  $x_D$  (m)

analisada, os intervalos entre 0,8 a 1; configuraram um padrão de nível estático entre 5 a 20 m e inferiores a 5 m. Ao centro, não há um padrão definido, de tal modo que, o nível estático apresenta-se ora entre o intervalo 5 a 20 m, ora superior a 20 m e ora inferior a 5 m. Ao Sudeste há um padrão que configurou o nível estático inferior a 5 m (Figura 9B).

Na análise de  $x_0$  e  $x_D$ (m), não se evidenciaram tendências, pois as retas de regressão apresentaram-se praticamente horizontais. Portanto, notou-se que é pouco provável que existam situações de deriva espacial em outras direções; posto que as retas não apresentaram declives consideráveis, para que se assumia a presença da deriva. Assim sendo, as hipóteses de estacionaridade não foram contrariadas, de tal modo que as estimativas de  $x_0$  e  $x_D$ (m), consideradas como variáveis regionalizadas, são válidas e coerentes, sem haver a necessidade de filtragem por meio de medida corretiva (e.g. krigagem de tendência).

### Variografia, ajustes e estimação geoestatística

#### Parâmetro: Ocorrência litológica – $x_0$

Para a distribuição espacial do parâmetro  $x_0$ , após várias tentativas experimentais, adotou-se as distâncias de deslocamento (*lag*) na direção  $x$  e  $y$  iguais a 560 e 990 m, respectivamente; e número de *lag* igual a 10, obtendo-se o mapa variográfico (Figura 10).

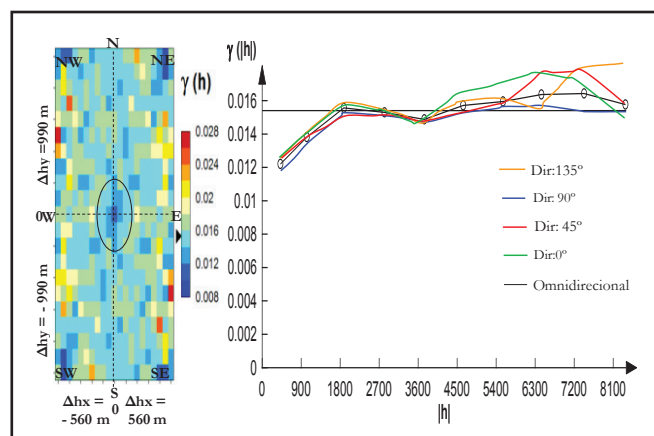


Figura 10 – Análise de deriva do parâmetro  $x_D$  (m)

Nessas condições, observou-se pouca anisotropia nas direções N-S e E-W; pois, dentro da elipse de busca, os pixels de tons em azul cuja tonalidade está abaixo da variância, foram predominantes no mapa variográfico. Considerou-se tolerância e incremento do *lag*, e *maximum bandwidth* iguais a 768 m, 930 m, 2.120 m; respectivamente.

Portanto, pelo variograma omnidirecional, verificou-se a não existência de tendência no conjunto de dados do parâmetro  $x_0$ , ratificado pela análise espacial preliminar de deriva. Notou-se que as amostras entraram num estado de aleatoriedade após 2.800m de distância  $b$  (Figura 10). Observou-se também que variogramas direcionais de 0°, 45°, 90° e 135°, tolerância angular igual a 45°, apresentaram similaridades, entrando em um estado de aleatoriedade após 2.100m de distância  $b$ .

Essa constatação foi ratificada pelo mapa variográfico que indicou pouca anisotropia nessas direções. Assim sendo, considerou-se a distribuição do referido parâmetro como isotrópico, sendo necessário ajuste ao variograma omnidirecional, apresentados pela Tabela 5.

Tabela 5 – Ajuste ao variograma omnidirecional

Parâmetros	IGF	$C_0$	$C$	$a$ (m)
SPH	1,60x10 <sup>-2</sup>	0,0108	0,0156	2.184
EXP	1,82x10 <sup>-2</sup>	0,0100	0,0158	2.688
GAU	1,65x10 <sup>-3</sup>	0,0118	0,0156	2.016

O melhor ajuste teórico foi associado ao modelo SHP. A escolha do variograma omnidirecional para esse processo justificou-se pelo mesmo ter um comportamento médio em relação aos variogramas direcionais.

Após melhor ajuste ao variograma ser relacionado ao



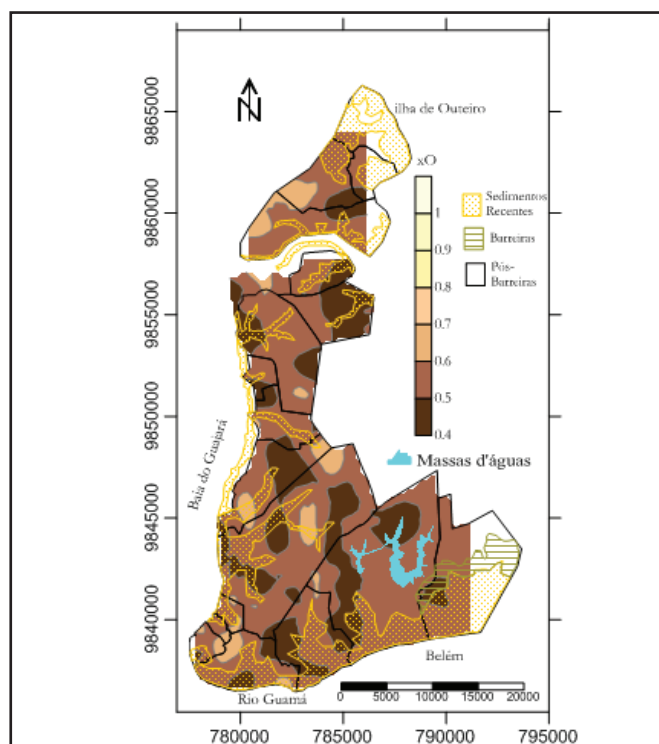


Figura 11 – Estimativa de  $x_0$  por modelo variográfico ajustado por GAU (0,0108; 0,0156; 2.184)

modelo GAU, foi possível estimar a variabilidade de  $x_0$ , ilustrado na Figura 11. O processo de validação do modelo variográfico revelou  $RMSE = 0,1273$  e  $MSE = 0,0162$ . Os erros agregados na estimativa da geoestatística por krigagem ordinária foram ilustrados pelo mapa do desvio-padrão (Figura 12).

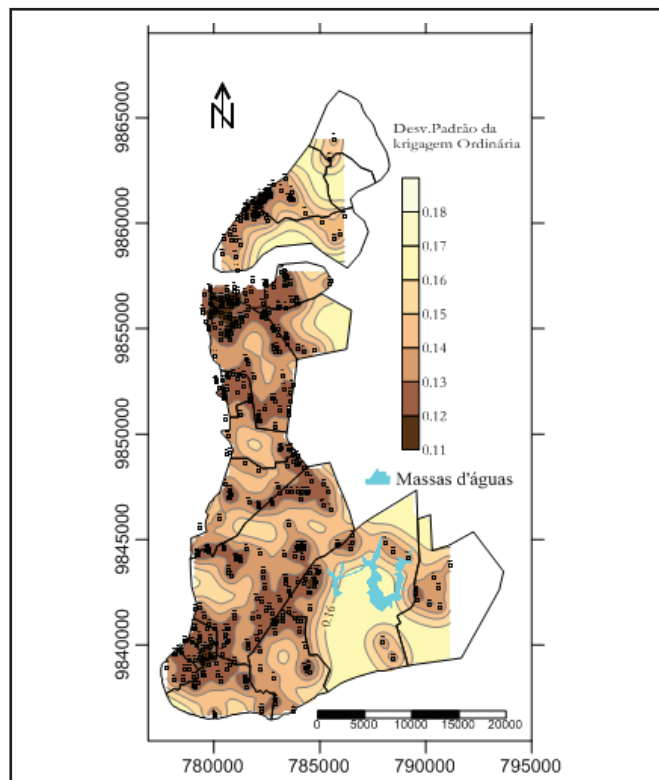


Figura 12 – Desvio-padrão da estimativa de krigagem ordinária de  $x_0$

Portanto, o mapa da ocorrência litológica revelou que o extrato da zona vadosa, em grande parte do domínio espacial foi mapeado como argiloso e argilo-arenoso (0,4 – 0,55). Essa estimativa foi associada a desvios-padrões entre 0,11 a 0,15. Desvio-padrões mais elevados, superiores a 0,15, estão relacionados a áreas de menor densidade de dados.

### Parâmetro: Nível estático – $x_D$ (m)

Para a distribuição espacial do parâmetro  $x_D$  foram considerados apenas os dados não outliers. Assim sendo, após várias tentativas experimentais, adotou-se a distância de  $lag$  na direção  $x$  e  $y$  iguais a 540 e 980 m, respectivamente; e um número de  $lag$  igual a 10; obtendo-se o mapa variográfico (Figura 13).

Nessas condições, observou-se anisotropias, de maior na direção N-W/S-E, e menor na direção N-E/S-W, para  $x_D$  (m). Chegou-se a tolerância e incremento do  $lag$ , e maximum bandwidth iguais a 550 m, 720 m, 3.700 m; respectivamente. Na análise do variograma omnidirecional verificou-se a não existência de tendência no conjunto de dado do parâmetro  $x_D$  (m), ratificado pela análise espacial preliminar de deriva. Pois, notou-se que as amostras entraram num estado de aleatoriedade após 3.500 m de distância entre pares de pontos ( $h$ ), descaracterizando a tendência da amostra (Figura 13).

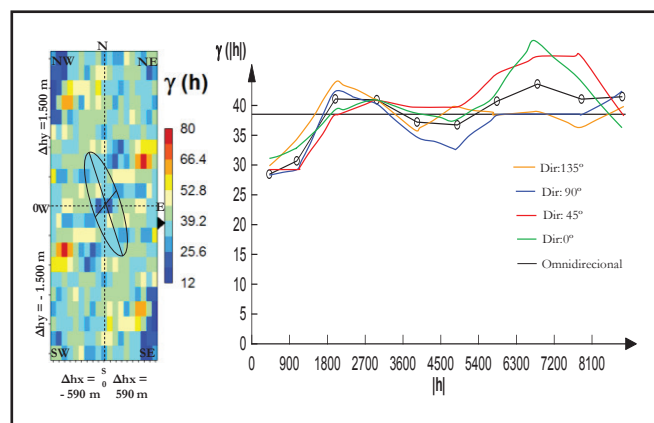


Figura 13 – Análise de deriva do parâmetro  $x_D$  (m)

Os variogramas direcionais de 45° e 135° foram associados à direção de menor e maior variabilidade, respectivamente; de TA igual a 45°. Observou-se que na direção de 45° as amostras entraram em um estado de aleatoriedade antes dos 3.500 m. Os ajustes ao variograma direcional estão apresentados pelas Tabelas 6. Obteve-se o melhor ajuste teórico pelo modelo GAU.

Tabela 6 – Ajuste ao variograma de direção 45° (TA: 45°)

Parâmetros	IGF	$C_0$	$C$	$a$ (m)
SPH	$1,81 \times 10^{-2}$	21,84	42,89	3.960
EXP	$1,81 \times 10^{-2}$	20,67	45,62	6.248
GAU	$1,47 \times 10^{-2}$	26,52	42,90	3.696

Na direção de 145°, as amostras entraram num estado de aleatoriedade a partir de 2.000m. Embora tenha sido testado por outros modelos teóricos, esse variograma foi melhor ajustado pelo modelo GAU (Tabelas 7).

Tabela 7 – Ajuste ao variograma de direção 135° (TA: 45°)

Parâmetros	IGF	C <sub>o</sub>	C	a (m)
SPH	1,93x10 <sup>-2</sup>	26,52	40,25	2.214,6
EXP	2,52x10 <sup>-2</sup>	24,56	40,86	2.640,0
GAU	1,85x10 <sup>-2</sup>	28,47	40,86	2.038,6

Portanto, dos valores acima foi possível escrever que o modelo variográfico direcional de 45° foi igual a:

$$\gamma_{45^\circ}(h) = 28,47 + 12,39 \left[ GAU \sqrt{\left(\frac{h_{45}}{\varepsilon}\right)^2 + \left(\frac{h_{135}}{2.038,6}\right)^2} \right] \quad (9)$$

E, o relativo à direção 135° foi expresso por:

$$\gamma_{135^\circ}(h) = 26,52 + 16,38 \left[ GAU \sqrt{\left(\frac{h_{45}}{3696}\right)^2 + \left(\frac{h_{135}}{\infty}\right)^2} \right] \quad (10)$$

onde GAU (h) é uma notação do modelo GAU normalizado, e (h) também foi normalizado em relação ao alcance a, pela seguinte expressão:

$$\left|\frac{h}{a}\right| = \sqrt{\left(\frac{h_{45}}{a}\right)^2 + \left(\frac{h_{135}}{a}\right)^2} \quad (11)$$

A modelagem da anisotropia combinada consistiu então em dividir o variograma da Figura 14 em quatro faixas, onde cada uma das faixas foi modelada como uma anisotropia geométrica.

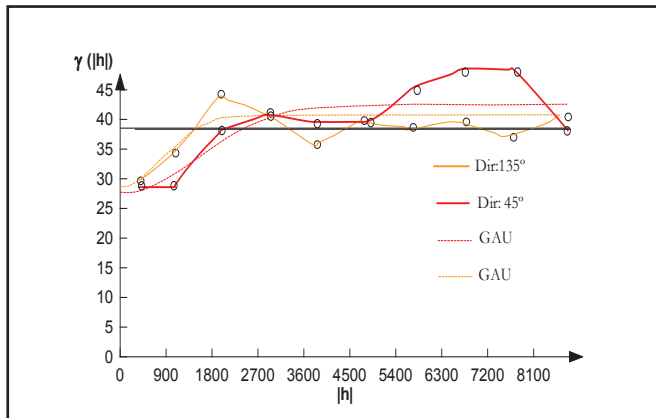


Figura 14 – Variograma direcional 45° e 135°, ambos ajustados por modelo GAU

Da Figura 15 conclui-se que a 1ª Faixa refere-se a C<sub>o</sub> = 28,47 (Equação 12). A anisotropia geométrica na 2ª Faixa foi estabelecida por um artifício que utilizou um modelo GAU com alcance muito pequeno (ε). Na direção 135°, notou-se que parte do modelo GAU participa com uma pequena contribuição igual a 1,95. Portanto, o modelo variográfico da 2ª Faixa foi expresso pela Equação (13).

$$\gamma_1(h) = 28,47 \quad (12)$$

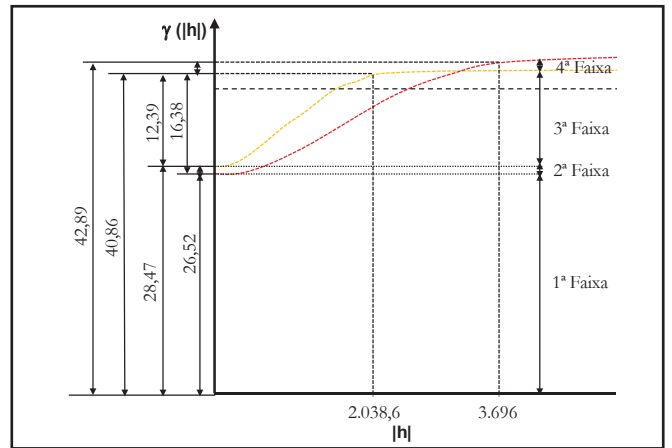


Figura 15 – Particionamento em faixas dos modelos variográficos ajustado para correção anisotrópica

$$\gamma_2(h) = 1,95 \cdot GAU \left[ \sqrt{\left(\frac{h_{45}}{\varepsilon}\right)^2 + \left(\frac{h_{135}}{2.038,6}\right)^2} \right] \quad (13)$$

Na 3ª Faixa, obteve-se um modelo variográfico pela direção 45° com alcance de 3.696 m; e, pela direção 135° e

$$\gamma_3(h) = 14,34 \cdot GAU \left[ \sqrt{\left(\frac{h_{45}}{3.696}\right)^2 + \left(\frac{h_{135}}{2.038,6}\right)^2} \right] \quad (14)$$

alcance de 2.038,6 m ambos por modelos GAU (Equação 14). Para estabelecer a anisotropia geométrica referente à 4ª Faixa empregou-se outro artifício, uma vez que notou-se que não existe modelo associado à direção 135°. Assim, atribui-se alcance muito grande (∞) para direção 135°. Como referiu Soares (2000), esse artifício foi utilizado apenas para estabelecer a anisotropia geométrica, não influenciando no modelo final. O modelo variográfico resultante da 4ª Faixa está expresso pela Equação (15).

$$\gamma_4(h) = 2,03 \cdot GAU \left[ \sqrt{\left(\frac{h_{45}}{3.696}\right)^2 + \left(\frac{h_{135}}{\infty}\right)^2} \right] \quad (15)$$

O modelo variográfico para qualquer distância e direção h, para estimar o valor de x<sub>D</sub>(m) foi determinado por: y(l) = y<sub>1</sub>(l) + y<sub>2</sub>(l) + y<sub>3</sub>(l) + y<sub>4</sub>(l).

Assim, o mapa do parâmetro x<sub>D</sub>(m), obtido por estimativa de krigagem ordinária pelo modelo variográfico, revelou que 46,36% do domínio espacial foi classificado como nível estático inferior a 5 m (0,9 – 1). Por outro lado, 53,14% da área estudada foi mapeada como nível estático entre e 5 a 20 m (0,8 – 0,9), e apenas 0,50% dessa área foi classificada como nível estático entre e 20 a 50 m (0,7 – 0,8) (Figura 16).

Os erros agregados na estimação por krigagem ordinária foram vinculados ao mapa do desvio-padrão, ilustrado na Figura 17. Essa estimativa foi associada a desvios-padrões entre 0,95 a 1,55.

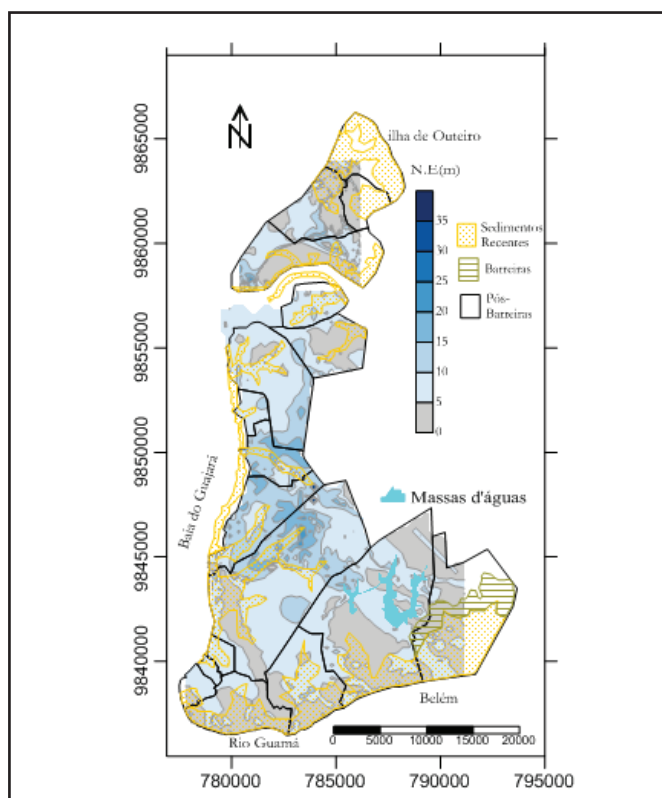


Figura 16 – Estimativa de  $x_D$  por modelo variográfico ajustado por GAU, após correção anisotrópica

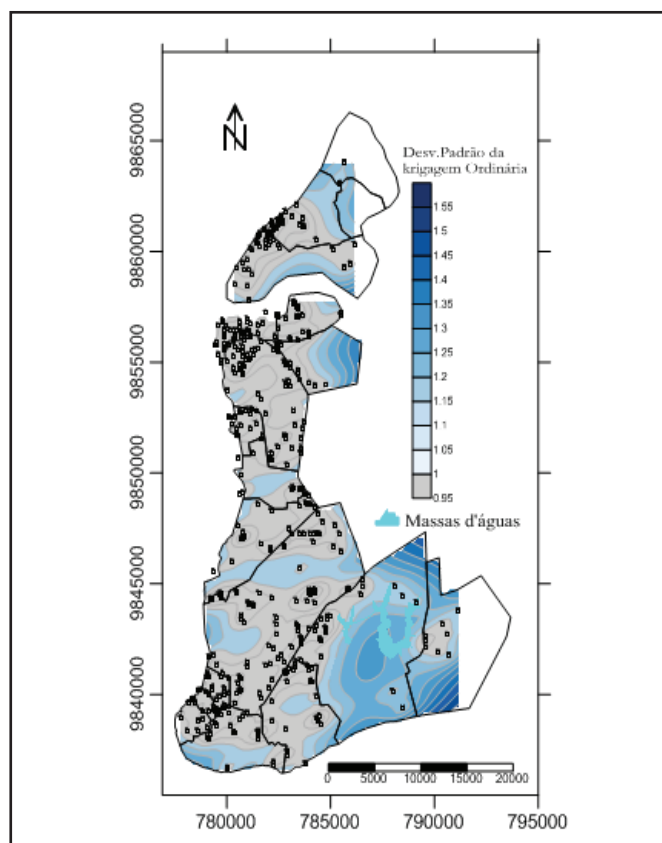


Figura 17 – Desvio-padrão da estimativa de krigagem ordinária de  $x_D$

Contudo, os desvio-padrões mais elevados estão relacionados a áreas em que não há uma relativa densidade de dados amostrais.

### Composição do mapa de vulnerabilidade intrínseca

A composição da vulnerabilidade de aquíferos freáticos foi feita a partir da interação multiplicativa das estimativas geoestatística de  $Z^*(x_0)$  e  $Z^*(x_D)$ . O resultado dessa interação é apresentado pela Figura 18.

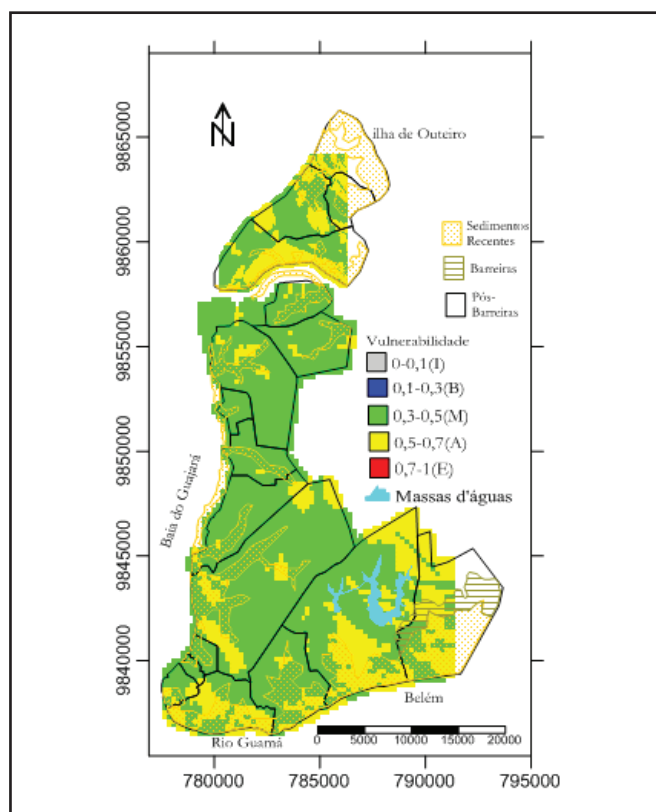


Figura 18 – Mapa de vulnerabilidade de aquíferos livres, pela interação multiplicativa de  $Z^*(x_0)$  e  $Z^*(x_D(m))$

Portanto, a partir dessa interação, a estimativa da vulnerabilidade intrínseca revelou que 62,31% do domínio espacial foram classificados como sendo de moderada vulnerabilidade (0,3 – 0,5). Em 37,69% da dessa área a vulnerabilidade foi classificada como alta (0,5 – 0,7).

### CONCLUSÕES

Esse trabalho desenvolveu uma metodologia para mapear a vulnerabilidade intrínseca de aquíferos livres por interação multiplicativa das estimativas dos parâmetros litologia e nível estático, obtidos por meio de estimativa geoestatística via krigagem ordinária.

O parâmetro relativo à ocorrência litológica foi estimado por meio de um modelo esférico. Contudo, não foram necessários ajustes anisotrópicos. O parâmetro nível estático foi estimado por um modelo gaussiano. Contudo, ajustes anisotrópicos entre variograma direcionais de menor e maior variabilidade espacial

foram procedidos. A estimação de ambos os parâmetros agregou pequenos erros, visualizados pelos mapas de desvio-padrão da krigagem. Logo, a interação das estimativas geoestatísticas desses parâmetros produziram o mapa de vulnerabilidade de aquíferos livres.

A metodologia aplicada é genérica a ponto de ser aplicada, novamente, em qualquer domínio espacial e hidrogeológico. Entretanto, ressalta-se que outros métodos podem ser inseridos na análise exploratória de dados e de deriva, bem como nas fases de variografia e estimação geoestatística. Ainda assim, se a metodologia proposta seja adotada, os resultados podem apresentar diferenças nas estimativas geoestatística, caso seja analisado apenas os valores do índice de vulnerabilidade. No entanto, a metodologia adotada apresentou-se como uma excelente estratégia para proteção das águas subterrâneas à contaminação.

## AGRADECIMENTOS

Aos revisores da revista RBRH. Ao CNPq/CsF, pela concessão de bolsa (Processo nº 246869/2012-7).

## REFERÊNCIAS

ALLER, L.; BENNET, T.; LEHR, H. J.; PETTY, J. R.; HACKETT, G. DRATIC: *A standardized system for evaluating groundwater pollution potential using hydrogeologic settings*. US. EPA-600/2-p. 87-035, 1987.

ASSAF, H.; SAADEH, M. Geostatistical assessment of groundwater nitrate contamination with reflection on DRATIC vulnerability assessment: the case of the Upper Litani Basin, Lebanon. *Water Resources Management*, v. 23, n. 4, p. 775–796, Mar. 2009.

BAALOUSHA, H. Assessment of a groundwater quality monitoring network using vulnerability mapping and geostatistics: a case study from Heretaunga Plains, New Zealand. *Agriculture Water Management*. v. 97, n. 2, p. 240–246, Feb. 2010.

CIVITA, M. *La carta della vulnerabilità degli acquiferi all'inquinamento*. Bologna: Pitagora Editrice, 1994.

CPRM (Companhia de Pesquisa de Recursos Minerais). *Projetos de Estudos Hidrogeológicos da Região Metropolitana de Belém e Adjacências*. Belém, 2001.

FERREIRA, G. T. *O uso de técnicas Geoestatísticas para o Mapeamento da Vulnerabilidade Natural do Aquífero guarani em área de recarga*. Aplicação na região de Ribeirão Bonito-SP. 2014. 97 f. Dissertação (Mestrado em Geociências e Meio Ambiente) - Universidade Estadual Paulista Júlio de Mesquita Filho, São Paulo, 2014.

FOSTER, S. S. D.; R. HIRATA. Groundwater pollution risk assessment: a methodology using available data. WHOPAHO/HPE-CEPIS Technical Manual. Lima, Peru, 1988.

LISBÔA, E.; BARP, A.; MENDES, R. Proposição Fuzzy Geoestatística para o Mapeamento da Vulnerabilidade Intrínseca de Aquíferos. *RBRH: revista brasileira de recursos hídricos*, v. 18, n. 1, p. 111-123, jan./mar. 2013.

MATTA, M. A. S. *Fundamentos hidrogeológicos para a gestão integrada dos recursos hídricos da Região de Belém/Ananindeua – Pará, Brasil*. 2002. 584 f. Tese (Doutorado em Hidrogeologia). Centro de Geociências, Universidade Federal do Pará, 2002.

NRC (National Research Council). *Groundwater vulnerability assessment: contamination potential under conditions of uncertainty*. Washington, D.C.: National Academy Press, 1993.

PALHETA, E. S. M. *Estudo da compartimentação e arcabouço neotectônico da ilha de Mosqueiro - Pará empregado no conhecimento hídrico subterrâneo*. 2008. 241 f. Tese (Doutorado em Geologia e Geoquímica). Instituto de Geociências, Programa de Pós-Graduação em Geologia e Geoquímica, Universidade Federal do Pará, Belém. 2008.

PANNATIER, Y. VARIOWIN: Software for Spatial Data Analysis in 2D. *Computational Statistics ; Data Analysis, Kelsterbach*, v. 25, n.2, p. 243-244, 1996.

RUSSO, A. S. *Estimativa da vulnerabilidade de aquíferos utilizando sistemas de informação geográfica e geoestatística-UGRHI-PCJ*. 2009. 85 f. Dissertação (Mestrado em Hidrogeologia e Meio Ambiente) - Instituto de Geociências, Universidade de São Paulo, São Paulo, 2009.

SOARES, A. *Geoestatística para as ciências da terra e do ambiente*. Lisboa: Instituto Superior Técnico (IST) Press, 2000.

VAN STEMPOORT, D.; EWERT, L.; WASSENAAR, L. Aquifer Vulnerability Index (AVI): a GIS compatible method for underground water vulnerability mapping. *Canadian Water Resources Journal*. v. 18, n.1, p. 25-37, 1993.

YAMAMOTO, J.K. Computation of Global Estimation Variance in Mineral Deposits. In: XIE, WANG; JIANG. *Computer Applications in the Minerals Industries*, 2001. p. 61-65.

## Contribuição dos autores:

Erico Gaspar Lisboa: Realização da coleta e tabulação dos dados secundários, construção e aplicação da metodologia, interpretação dos resultados e fechamento do texto.

Jorge Manuel Machado De Carvalho: Revisão da metodologia proposta e interpretação dos resultados geoestatísticos, e fechamento do texto.

Ronaldo Lopes Rodrigues Mendes: Interpretação das análises geológicas e hidrogeológicas, e fechamento do texto.