

# Comportamento espacial da precipitação média pluviométrica do estado de Minas Gerais

Adrielle Giaretta Biase <sup>1</sup>

Eder Comunello <sup>2</sup>

Murilo Vianna <sup>3</sup>

Paulo Justiniano Ribeiro Júnior <sup>4</sup>

## Resumo

A chuva pode ser considerada a principal forma de suprimento hídrico para as atividades humanas e econômicas, principalmente na agricultura. O objetivo deste trabalho foi realizar o zoneamento climático, com a variável precipitação pluvial, no Estado de Minas Gerais, com o auxílio da ferramenta geoestatística. O estudo foi desenvolvido utilizando dados de precipitação pluvial média anual de 396 estações climáticas do estado de Minas Gerais, referente a média do período de 30 anos de dados, 1980 a 2010. Toda análise foi computada no *software* R no pacote da *GeoR*. Nesse estudo possibilitou conhecer a espacialização da distribuição média anual das chuvas na região de MG, assim como identificar os locais em que a população do polígono das secas mais sofre com a falta de uniformidade da precipitação.

Palavras-chave: Geoestatística, precipitação, Minas Gerais e agricultura de precisão.

## 1 Introdução

Fenômenos naturais como precipitação, temperatura, deficiência hídrica, constitui-se um grande objeto de estudo a ser pesquisado principalmente na área agrícola e em projetos de engenharia. Dentre os parâmetros meteorológicos de estudo, a precipitação é a que apresenta a maior variabilidade temporal e espacial sendo, portanto, o de maior dificuldade na estimação. Daí o interesse em aprofundar melhor o estudo dessa variável no Estado de Minas Gerais.

A chuva pode ser considerada a principal forma de suprimento hídrico para as atividades humanas e econômicas, principalmente na agricultura, setor em que a precipitação é usada, direta ou indiretamente, em processos de tomada de decisão, como segue: no manejo da irrigação, no risco de incêndios, na previsão de geadas, no zoneamento agroclimático, na ocorrência de pragas e doenças, na classificação climática, dentre outros segundo Almeida et al. (2011).

Usinas hidrelétricas de acumulação dependem muito da precipitação na região. Uma região seca pode ter um impacto grandes sobre a produção de eletricidade. A determinação de boas estimativas de variáveis físicas é o ponto base e mais importante para o desenvolvimento de projetos dessa engenharia. Estruturas hidráulicas destinadas à contenção do excesso de chuva, como barragens, bacias de contenção, terraços e outras, são projetadas mediante o uso de um bom estimador para uma precipitação que satisfaça a segurança desejada e promova um projeto economicamente racional, Mello et.al (2003).

---

<sup>1</sup>ESALQ/USP - e-mail: [adrielegbiase@usp.br](mailto:adrielegbiase@usp.br)

<sup>2</sup>ESALQ/USP - e-mail: [ecomunel@gmail.com](mailto:ecomunel@gmail.com)

<sup>3</sup>ESALQ/USP - e-mail: [murilodsv@gmail.com](mailto:murilodsv@gmail.com)

<sup>4</sup>UFPR - e-mail: [paulojus@c3sl.ufpr.br](mailto:paulojus@c3sl.ufpr.br)

O objetivo deste trabalho foi realizar o zoneamento climático, com a variável precipitação pluvial, no Estado de Minas Gerais, com o auxílio da ferramenta geoestatística, a fim de fornecer suporte à agricultura de precisão, aplicações agro-ambientais e projetos de engenharia.

## 2 Material e Métodos

### 2.1 Descrição dos dados

Os dados utilizados nos processos são pontos georreferenciados referentes a média de 30 anos relativos aos anos de 1980 até 2010, das estações meteorológicas do Estado de Minas Gerais localizado na Região Sudeste do Brasil, entre os paralelos 1413' e 2255' de latitude Sul e os meridianos 3951' e 5102' de longitude Oeste. A variável referente ao estudo é a precipitação média anual, provenientes de 394 estações meteorológicas no estado de Minas Gerais, obtida acumulando-se a precipitação de cada ano e obtendo a média dos anos acumulados em milímetros. Os dados foram obtidos pelo Sistema de Monitoramento Agrometeorológico (Agritempo), em que é possível obter informações meteorológicas e agrometeorológicas de diversos municípios e estados brasileiros, pelo site: <[www.agritempo.gov.br](http://www.agritempo.gov.br)>.

Para todo o desenvolvimento e execução do trabalho optou-se utilizar o programa R versão 2.15.1 (R Development Core Team, 2012) e o pacote *geoR*. Os dados da precipitação média anual foram organizados de acordo com o georeferenciamento das coordenadas (longitude e latitude), respectivamente na unidade *utm*. Foi considerada uma covariável, a altitude, para o estudo da dependência espacial.

Realizou-se a análise exploratória não espacial na variável precipitação, essa análise preliminar auxilia apontando alguns possíveis indicativos das características dos dados, como *outlier* e o ajuste da distribuição normal, Diggle e Ribeiro (2007). Após essa tarefa foram aplicadas funções do pacote *geoR* para análises geoestatísticas. Foram usadas as funções *variog*, *eyefit*, *variogfit*, *likfit*, *variog4*, *krige.conv* e *xvalid*. A função *eyefit* foi usada com a finalidade de gerar chutes iniciais para ajustar um modelo teórico ao semivariograma na função *variogfit*. Os parâmetros estimados pela função *variogfit* foram usados na função *likfit*.

Para a seleção de modelos primeiramente foi feito o teste da Razão de Verossimilhança que é baseado na razão entre as duas verossimilhanças dos modelos mais simples,  $l(\tilde{\theta})$ , e o modelo mais complexo,  $l(\hat{\theta})$ . A deviance do teste é apresentada sendo

$$D = -2[\log(l(\tilde{\theta})) - \log(l(\hat{\theta}))] \sim \chi_r^2 \quad (1)$$

que é assintoticamente distribuída como uma qui-quadrado com  $r$  graus de liberdade e serve para testar a hipótese  $H_0$ : o modelo mais simples é adequado. Quando essa hipótese for rejeitada, conclui-se que o modelo mais complexo (com maior número de parâmetros) é adequado. Se o modelo mais simples for adequado, os valores da função de verossimilhança avaliadas em  $\hat{\theta}$  e  $\tilde{\theta}$  devem estar próximos, indicando que os dados estão dando suporte ao modelo com menor número de parâmetros, Mariguele et al. (2011).

Embora o método seja eficiente e bastante usado em trabalhos científicos, é limitado pelo fato de poder testar apenas dois modelos de cada vez, sendo que um deles deve ser um caso especial do outro (modelos encaixados). Num segundo passo, para comparar modelos não encaixados, usou-se o Critério de Informação de Akaike (AIC).

Foi realizada a validação cruzada para confirmar a escolha do melhor modelo, para isso foi usada a

função *xvalid*. O procedimento de validação cruzada consiste na remoção de um valor determinado em campo com posterior estimativa deste ponto, considerando os parâmetros do modelo ajustado (Teixeira, 2011).

Após selecionar o modelo usando o teste da Razão de Verossimilhança e o Critério de Informação de Akaike (AIC) foi realizado o estudo de anisotropia. Para identificar a presença de anisotropia e avaliar a dependência espacial da variável precipitação média, foi plotado o semivariograma nas direções convencionais de 0°, 45°, 90° e 135°, usando a função *variog4* com o ângulo de tolerância de  $\pi/8$ , ou seja, o próprio *default* da função, não sendo impedido de trabalhar com outras direções. Com o desenvolvimento desse semivariograma foi possível obter informações sobre a existência ou não de tendência direcional, o efeito proporcional e a continuidade espacial da variável precipitação média anual no estado de MG. Se o gráfico apresentar variação gradativa numa dada região ou direção indicará possível existência de uma tendência direcional e, conseqüentemente, a falta de estacionariedade (Libardi et al., 1996).

Ao detectar a presença de isotropia foi construído o mapa temático de variabilidade espacial para precipitação média anual na área do estado de MG, utilizando a técnica de interpolação por krigagem ordinária para pontos não-amostrados, usando a função *krige.conv*. A descrição e comandos utilizados podem ser encontrados no apêndice A.

### 3 Resultados e discussões

A fundamentação da estatística inferencial é calcada na suposição de normalidade e/ou na extensão do teorema central do limite em sua maioria. No entanto, a normalidade não é uma exigência da geostatística, mas é conveniente que a distribuição não apresente caudas muito alongadas, pois as estimativas da krigagem podem ser comprometidas, já que elas são baseadas nos valores médios (ISAACS e SRIVASTAVA, 1989). Foram apresentados o Box-Plot e o histograma nas Figuras (1) e (2), respectivamente, como forma de análise descritiva da variável precipitação total. Notou-se ausência de *outliers* pelo Box-Plot e o desvio de simetria foi um pouco mais marcante visto pelo histograma.

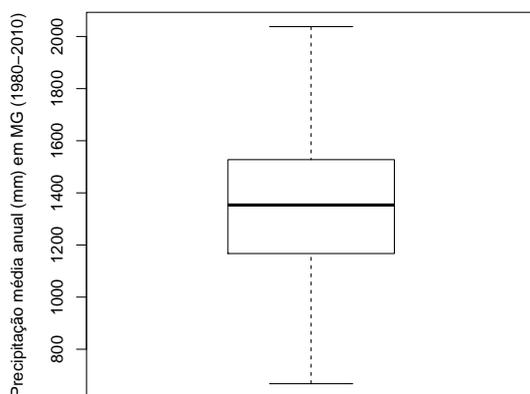


Figura 1 Box Plot: Precipitação média anual (mm) em MG (1980-2010)

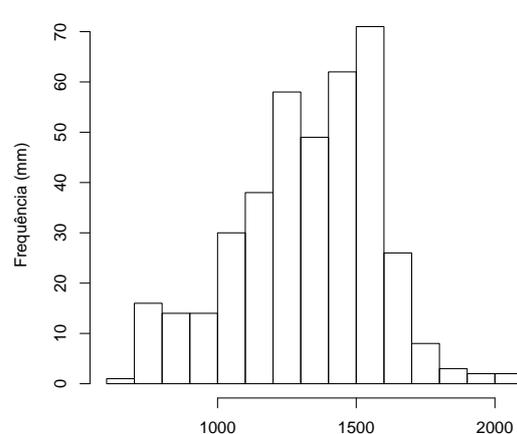


Figura 2 Histograma: Precipitação média anual (mm) em MG (1980-2010)

Na Tabela (1) são apresentadas as estatísticas descritivas da variável precipitação média e da covariável altitude. Na Tabela (2) são apresentadas as análises descritivas da localização da mínima e da máxima coordenada (longitude e latitude) das estações climatológicas do Estado de Minas Gerais, e a menor e maior distância entre as estações. O coeficiente de variação dos dados foi 19.3821%.

Tabela 1 Estatísticas descritiva da precipitação total mensal (*mm*), temperatura (°C) e da deficiência hídrica (*mm*) das estações climatológicas do Estado de Minas Gerais.

		Estatísticas descritivas						
		Média	D.P.	min	1 quantil	mediana	3 quantil	max
variável	precipitação	1326.0	256.92	667.70	1168.00	1353.00	1527.00	2038.00
covariável	altitude	743.40	256.70	83.00	575.50	766.00	908.50	1389.00

Nota: D.P.: desvio padrão, min.: mínimo e max.: máximo.

Tabela 2 Estatísticas descritiva das coordenadas das estações climatológicas do Estado de Minas Gerais.

coordenadas	mínima (utm)	máxima (utm)
longitude	$15.5566 \times 10^4$	$120.5160 \times 10^4$
latitude	$748.1482 \times 10^4$	$842.2257 \times 10^4$
Entre as estações	distância mínima (utm) $0.1046 \times 10^4$	distância máxima (utm) $112.3133 \times 10^4$

### 3.1 Análise Exploratória Espacial

Diggle e Ribeiro Jr (2007) definem como primeiro estágio em uma análise espacial a investigação gráfica da variável resposta em relação às suas coordenadas. Essa inspeção possibilita detectar *outliers* e tendências espaciais. O gráfico obtido para a variável precipitação média anual é apresentado na Figura (3). A coordenada *x* refere-se à longitude e a coordenada *y* à latitude. Nota-se o comportamento semelhante entre os pontos mais próximos. A precipitação média anual (*mm*) é maior na região Sul e no oeste de MG (Triângulo Mineiro). Por outro, é possível verificar a presença de alguns pontos diferentes de seus vizinhos, indicando a presença de *outliers* espaciais, é o caso da estação no município de Abadia dos Dourados com a precipitação média anual de 1101.8 *mm*, abaixo das precipitações médias dos vizinhos próximos. As estações próximas a esse município apresentam precipitações acima da mediana das 394 estações estudadas em MG, é o caso dos município de Pântano (1714.6 *mm*), Monte Carmelo (1522.93 *mm*) e Cascalho Rico (1713.02 *mm*).

Vale destacar também a detecção de tendência espacial da precipitação em MG, observa-se que os menores valores da precipitação estão associados com altos valores da coordenada *y* (latitude), essa mesma relação é menos perceptível ao longo da coordenada *x*. Esse tipo de tendência pode ser claramente detectado em gráficos de dispersão da variável resposta contra cada uma das coordenadas espaciais. Nas Figuras (4) e (5) é apresentado os gráficos de dispersão da variável precipitação com as coordenadas *x* e *y*, respectivamente.

Com a finalidade de estudar o comportamento espacial foram levados em consideração o estudo da tendência com as coordenadas, longitude e latitude, e a covariável altitude. Foram considerados cinco tipos de tendência espacial, combinando algebricamente a longitude, latitude e altitude. As tendências consideradas foram para o primeiro modelo a longitude e a altitude ( $x + y$ ). No segundo modelo foi considerada a combinação linear da longitude, latitude e a altitude ( $x+y+z$ ). Também foram considerados

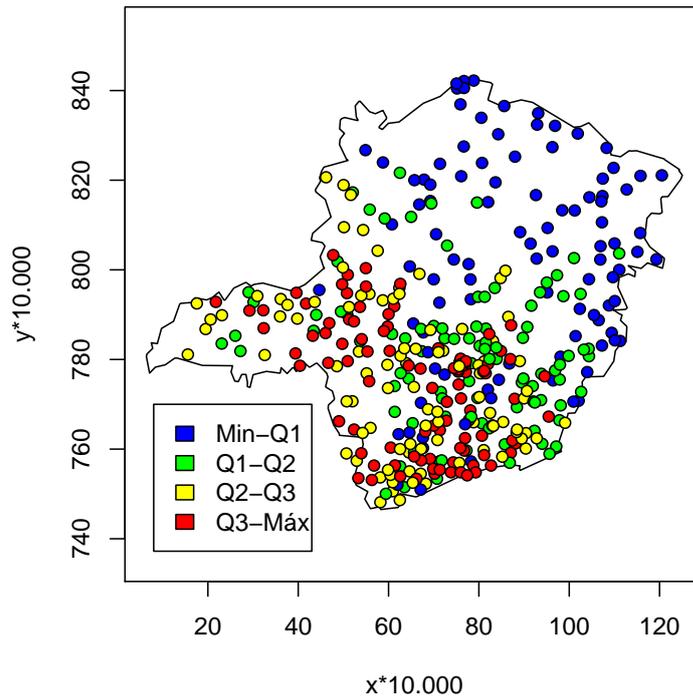


Figura 3 Gráfico de círculos representando os pontos amostrados em suas coordenadas específicas, considerando a variável precipitação média anual, dividido nos quantis.

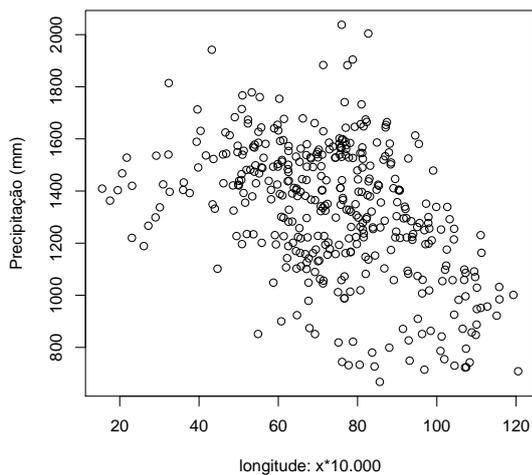


Figura 4 Relação entre precipitação e a coordenada geográfica x.

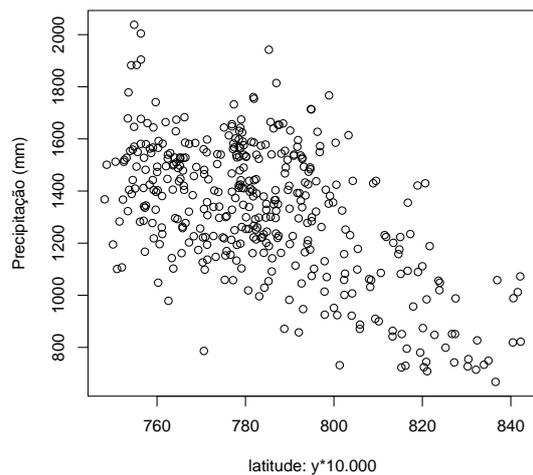


Figura 5 Relação entre precipitação e a coordenada geográfica y.

três modelos com tendências quadráticas representados por  $x + y + x^2 + y^2$ ,  $x + y + xy + x^2 + y^2$  e  $x + y + x^2 + y^2 + z$ . A análise gráfica das dispersões e da distribuição de probabilidade residuais foram apresentados

nas Figuras (6), (7), (8), (9) e (10), respectivamente, para cada modelo considerado. Notou-se que os modelos quadráticos modelaram melhor a distribuição residual. Como discutido por Diggle e Ribeiro Jr (2007), idealmente o modelo de uma tendência deve ter uma interpretação física natural. Assim a escolha de um modelo simples e que explique a maior parte variabilidade espacial seria o ideal. Modelos mais complexos são de modo geral mais difíceis de interpretar.

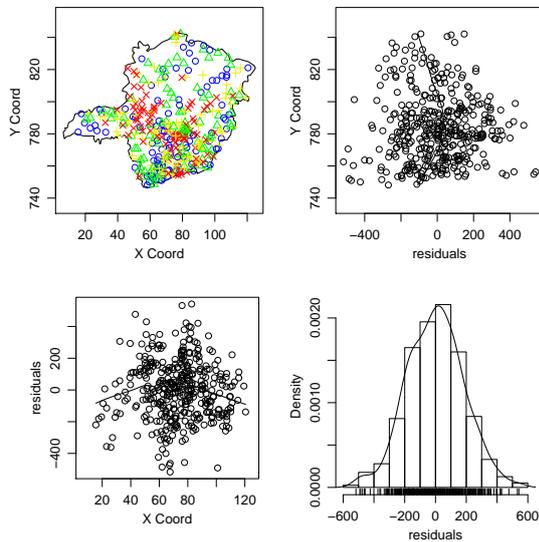


Figura 6 Análise gráfica residual considerando a tendência, covariável altitude ( $x+y$ ).

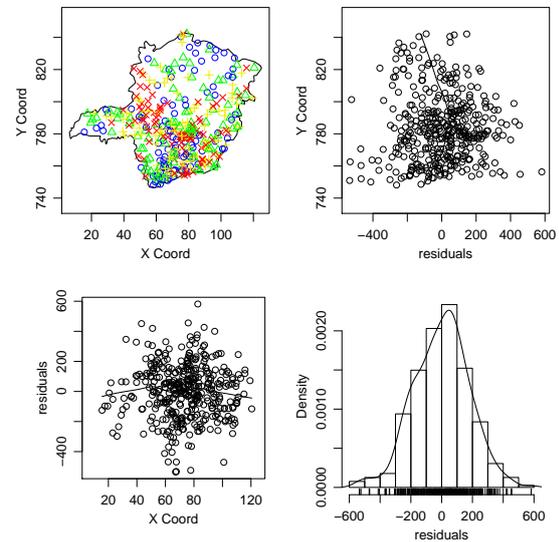


Figura 7 Análise gráfica residual considerando a tendência longitude e latitude ( $x+y+z$ ).

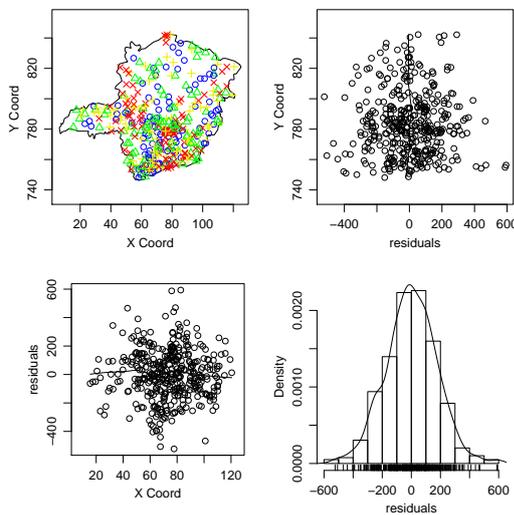


Figura 8 Análise gráfica residual considerando a tendência longitude, latitude e altitude ( $x + y + x^2 + y^2$ ).

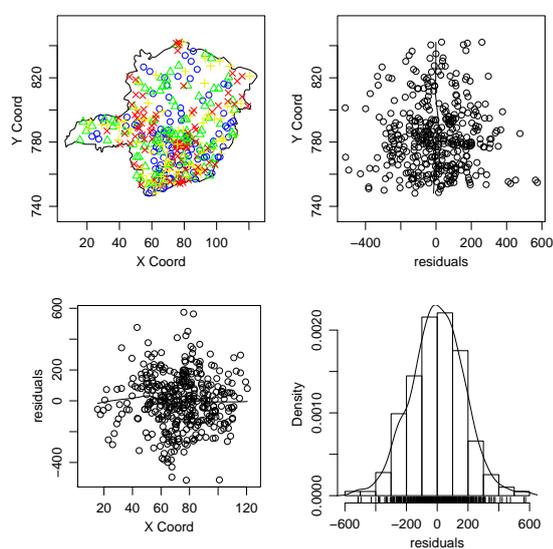


Figura 9 Análise gráfica residual com inserção de tendência ( $x + y + xy + x^2 + y^2$ ).

Foram realizados os variograma empíricos, que é a representação gráfica das variâncias das diferenças de distâncias entre todos os pares de pontos ( $v_{ij}$ ) contra as distâncias correspondentes. Espera-se que, com o aumento da distância entre os pontos, a correlação entre eles aproxime-se de zero e a esperança de  $v_{ij}$  aproxime-se de  $\sigma^2$ . Essa análise é uma ferramenta útil na análise exploratória espacial e é conhe-

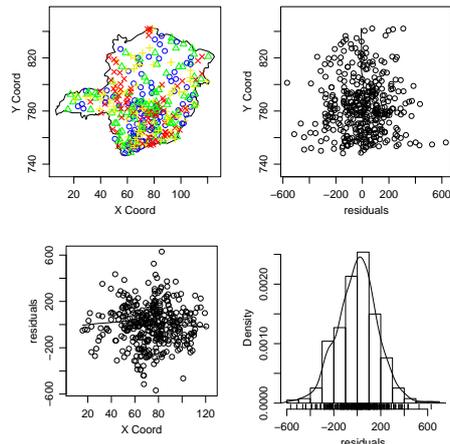
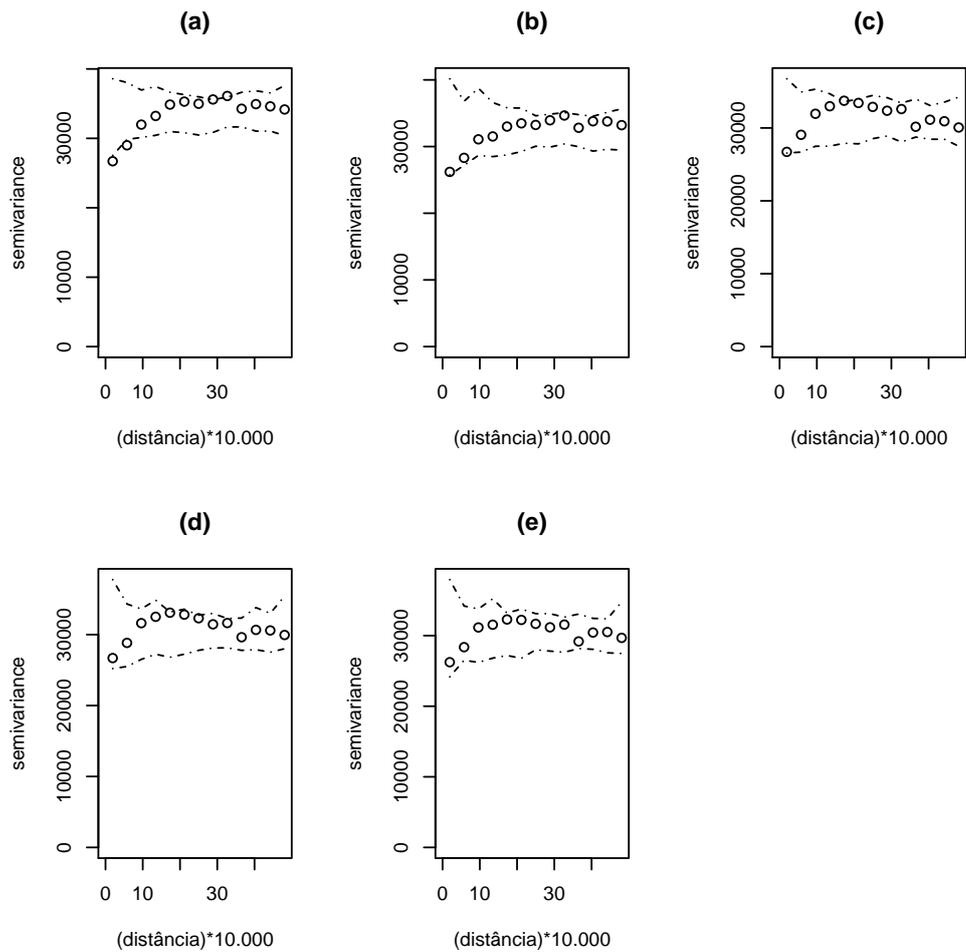


Figura 10 Análise gráfica residual considerando a tendência  $(x + y + x^2 + y^2 + z)$ .

cida com os ajustes dos semi-variogramas. Foram realizados os ajustes dos semivariogramas dos cinco modelos em que foram consideradas as tendências mencionadas anteriormente, na Figura (11).



Nota: **(a)**  $x+y$ , **(b)**:  $x+y+z$ , **(c)**:  $x + y + x^2 + y^2$ , **(d)**:  $x + y + x^2 + y^2 + xy$  e **(e)**:  $x + y + x^2 + y^2 + z$

Figura 11 Ajuste dos semi-variogramas considerando diferentes tendências

Os variogramas são usados apenas como uma ferramenta de análise exploratória e não como base para uma inferência formal, Diggle e Ribeiro Jr (2007). Para essas inferências será utilizado o método da máxima verossimilhança, cujos resultados serão apresentados na subseção (3.2).

Para todos modelos ajustado houve uma estabilização a longa distância nos semivariogramas (Figura 11). Observa-se também que as semivariâncias tiveram uma queda aparente em distâncias mais longas. Entretanto, utilizou-se da técnica do envelope de variogramas empíricos computados de permutações aleatórias. Notou-se indícios de não haver dependência espacial quando computados os variogramas cujos os modelos eram quadráticos. Como discutido anteriormente, resultados mais consistentes serão obtidos a seguir.

### 3.2 Estimação dos parâmetros e predição espacial

Para estimação dos parâmetros do modelo geoestatístico optou-se por usar o método da máxima verossimilhança, pois possuem certas propriedades ótimas, sob condições de regularidade. Numa análise preliminar (resultados não mostrados) constatou-se que o modelo Esférico foi o que melhor descreveu o conjunto de dados, com base no Critério de Informação de Akaike (AIC). Assim, para a variável resposta precipitação assumiu-se o modelo Esférico estimando os parâmetros e considerando cada tendência. Para a tendência quadrática,  $x + y + x^2 + y^2 + z$ , constatou-se a ausência de dependência espacial. As estimativas dos parâmetros são apresentadas na Tabela 3.

Tabela 3 Estimativas obtidas por máxima verossimilhança para os parâmetros  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$ ,  $\beta_4$  e  $\beta_5$ , efeito-pepita ( $\tau^2$ ), variância ( $\sigma^2$ ) e alcance ( $\phi$ ), considerando as variável precipitação média anual (1980-2010) em MG.

modelo	Tendências	Parâmetros estimados								
		$\beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\sigma^2$	$\phi$	$\tau^2$
1	$x + y$	7083	-4.9938	-6.9031	0.0000	0.0000	0.0000	10204	$21.64 \times 10^4$	23056
2	$x + y + z$	6660.23	-4.5717	-6.5131	0.1244	0.0000	0.0000	9086	$21.64 \times 10^4$	23109
3	$x + y + x^2 + y^2$	-36755.23	3.1455	102.8629	-0.0564	-0.0691	0.0000	8093	$17.79 \times 10^4$	23079
4	$x + y + x^2 + y^2 + xy$	-32953.17	44.28	89.3086	-0.0445	-0.0579	-0.0540	7645	$17.5 \times 10^4$	23165
5	$x + y + x^2 + y^2 + z$	-49287.9	1.66	133.7688	-0.0406	-0.0882	0.1491	0	0	0

Nota: x: longitude, y:latitude e z: altitude,

Realizou-se a análise dos resíduos testando a normalidade do resíduo com base no teste de normalidade de Shapiro-Wilk, não foi necessário usar a transformação de Box-Cox, os resíduos possuíam uma boa aproximação de uma distribuição de probabilidade normal. Os resultados da estatística e o Valor P são apresentados na Tabela 4.

Tabela 4 Teste de normalidade Shapiro-Wilk para os resíduos dos modelos considerando a tendência para a variável precipitação

	Modelo esférico	Teste de normalidade Shapiro-Wilk	
	Tendência	W	Valor P
Precipitação	$x + y$	0.9950	0.2357
	$x + y + z$	0.9943	0.1449
	$x + y + x^2 + y^2$	0.9956	0.3282
	$x + y + x^2 + y^2 + xy$	0.9952	0.2570
	$x + y + x^2 + y^2 + z$	0.9958	0.3797

Nota: z: altitude, x: longitude e y:latitude

Para escolher o melhor modelo primeiramente foi realizada o teste da Razão de Verossimilhança para os casos em que os modelos eram encaixados e posteriormente usou-se o Critério de Informação de Akaike (AIC) para os modelos não encaixados. Os resultados da estatística do teste da Razão de Verossimilhança é apresentada na Tabela 5.

Tabela 5 Teste de normalidade da Razão de Verossimilhança para os modelos encaixados

	$H_0$ : Modelo simples é adequado			Teste da Razão de Verossimilhança	
	simples	vesus	complexo	$D$	$\chi_r^2$
Precipitação	$x + y$	vesus	$x + y + z$	5,7100	0,0160
	$x + y$	vesus	$x + y + x^2 + y^2$	1,3497	0,0623
	$x + y + x^2 + y^2$	vesus	$x + y + x^2 + y^2 + xy$	1,3497	0,2453

Nota: z: altitude, x: longitude e y: latitude

Para os três testes realizados na tabela, a hipótese  $H_0$  foi rejeitada, ou seja, os modelos mais complexos foram o que explicaram melhor a variabilidade espacial no estado de MG. Entre a escolha do modelo  $x + y + z$  e  $x + y + x^2 + y^2 + xy$  foi usado o método do AIC, cujos os resultados foram respectivamente, 5169 e 5172, concluindo que o modelo com menor número de parâmetros foi mais adequado. Ou seja, o modelo esférico que leva em consideração a tendência linear da longitude, latitude e altitude é o que melhor explica a variabilidade espacial da precipitação média anual em MG.

Pacheco, Silva Jr e Oliveira (2012) realizaram um estudo entre algumas estações climatológica de MG, para verificar a correlação entre precipitação e altitude. As cidades consideradas no estudo foram Barbacena, Juiz de Fora, Lavras, Caparaó e Viçosa. Nesse, estudo notou-se uma baixa relação entre precipitação e altitude. Porém os autores mencionam a necessidade de adensamento da rede de monitoramento para melhorar a qualidade dos dados e permitir uma caracterização climática mais pontual, pois parte do banco de dados utilizados para a pesquisa continha falhas.

Porém, num estudo mais completo usando a metodologia de análise de modelos digitais de elevação em microbacias, Valeriano e Picini (2003) corrobora com os resultados apresentados nesse estudo e mostram que a variação da precipitação anual acompanha os desníveis altimétricos, sendo, portanto, a altitude influente na distribuição da precipitação.

O estudo e a análise estrutural do fenômeno por meio da geoestatística também devem considerar que a estrutura da variabilidade espacial está relacionada com a estrutura natural do fenômeno. Assim, algumas características devem ser verificadas, como a continuidade espacial e a ausência de anisotropia, Ortiz (2002). Uma variável é isotrópica quando o padrão de continuidade espacial é idêntico em todas as direções. A isotropia é de suma importância para a estimativa de locais não-amostrados (Guedes et al., 2008).

Com base nas estimativas do modelo esférico com tendência linear,  $x + y + z$ , obtidas por meio da função *likfit*, foi realizado o estudo para verificar se havia presença de anisotropia, o gráfico para os semivariogramas direcionais para a precipitação média anual é apresentado na Figura 12.

Observa-se que o padrão de continuidade espacial é semelhante em todas as direções consideradas, o que proporciona uma melhor acurácia do mapa de variabilidade espacial a serem gerados por meio da krigagem. Assim, uma vez verificada a ausência da anisotropia foi possível construir o mapa que expressa a variabilidade da precipitação avaliada na área em estudo, representando uma superfície contínua que caracteriza o comportamento da variável no estado de MG. Primeiramente é apresentado o gride, em que foram colocados 11.543 pontos ( $97 \times 119$ ), considerando uma malha quadrangular igualmente espaçada com distância de 10 km entre os pontos, Figura 13. O resultado da krigagem com os índices

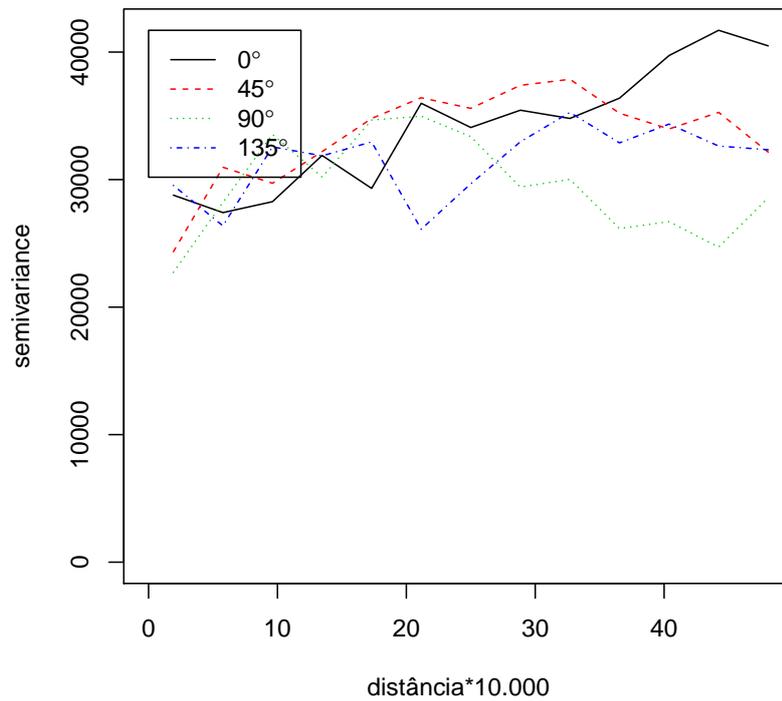


Figura 12 Semivariogramas direcionais para a precipitação média anual - MG

da precipitação pluviométrica média anual de Minas Gerais é a apresentado na Figura 14. Nesse mapa percebe-se claramente que ao longo das regiões sul e centro-oeste os índices pluviométricos são mais elevados.

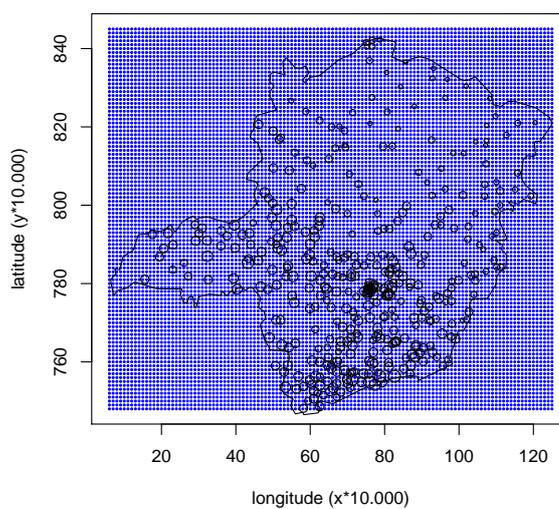


Figura 13 Grid para predição - precipitação média anual - MG

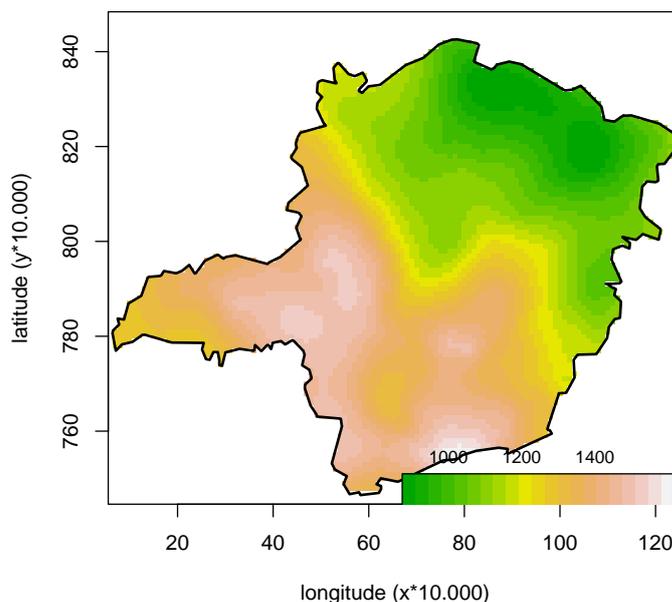


Figura 14 krigagem - precipitação média anual - MG

## 4 Conclusões

O conhecimento da distribuição média anual das chuvas na região possibilitou conhecer sua espacialização, assim como identificar os locais em que a população do polígono das secas mais sofre com a falta de uniformidade da precipitação. A Geoestatística pode ajudar nas análises da rede meteorológica e também em seu planejamento.

O Território Médio Jequitinhonha (TMJ), localizado no Vale do Jequitinhonha, porção norte do estado de Minas Gerais, é fortemente marcado por condições climáticas adversas, tendendo para a semi-aridez, com precipitações anuais abaixo de 1.000 milímetros, que aliadas à condição de fertilidade dos solos, relativamente baixa para as principais culturas agrícolas, tornam a sobrevivência difícil, especialmente no meio rural.

O zoneamento agrícola de risco climático vem sendo gradativamente ampliado e utilizado em larga escala no Brasil. O ordenamento territorial da precipitação média anual no estado de MG buscou auxiliar o incentivo de políticas agrícolas, a seguridade agrícola (instituições financiadoras de crédito rural), o planejamento agrícola e o de engenharias no estado de MG.

## 5 Continuidade do trabalho

O próximo passo seria utilizar métodos multivariados de análise de agrupamento e tentar classificar os grupos de regiões com climas semelhantes. E comparar com as probabilidades interpoladas na região do estado de Minas Gerais por meio da Geoestatística.

## 6 Agradecimentos

Agradeço aos colegas da turma de Geoestatística (ESALQ/USP), 2012, que muito contribuíram com esse trabalho com trocas de experiência e ideias. Em especial a Naimara, Djair e Gabriel.

## Referências

- [1] ALMEIDA, A. Q. et al. Geoestatística no estudo de modelagem temporal da precipitação. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v.15, n.4, p. 354-358, Jan. 2011.
- [2] DIGGLE, P. J.; RIBEIRO JR, P. J., **Model-based Geostatistics**. Londres: Springer Series in Statistics, 2007, 230 p.
- [3] GUEDES, L. P. C. et al. Anisotropia no estudo da variabilidade espacial de algumas variáveis químicas do solo. **R. Bras. Ci. Solo**, Viçosa, v.32, n. 1, p. 2217-2226, 2008.
- [4] ISAAKS, E.H.; SRIVASTAVA, R.M. An introduction to applied geoestatistics. New York: Oxford University, p. 561, 1989.
- [5] LIBARDI, P.L.; MANFRON, P.A.; MORAES, S.O. TOUN, R.L. Variabilidade da umidade gravimétrica de um solo hidromórfico. **R. Bras. Ci. Solo**, Viçosa, v.20, n. 1, p. 1-12, 1996.
- [6] MARIGUELE, K. H. et al. Métodos de análise de dados longitudinais para o melhoramento genético da pinha. **Pesq. agropec. bras.**, Brasília, v.46, n.12, p.1657-1664, dez. 2011.
- [7] MELLO, K. H. et al. Modelos matemáticos para predição da chuva de projeto para regiões do Estado de Minas Gerais. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.7, n.1, p.121-128, fev. 2003
- [8] ORTIZ, G. C. **Aplicação de métodos geoestatísticos para identificar a magnitude e a estrutura da variabilidade especial de variáveis físicas do solo**. 2002, v.1. 75 p.
- [9] R DEVELOPMENT CORE TEAM. R: A Language and Environment for Statistical Computing. **R Foundation for Statistical Computing**, 2012. Disponível em: <<http://www.r-project.org>>. Acesso em: 14 setembro 2012.
- [10] RIBEIRO JR, P. J.; DIGGLE, P. J. , geoR: A package for geoestatistical analysis. **R-NEWS**, v. 1, n.2, p. 15-18, 2001
- [11] TEIXEIRA, G. C. et al. Estudo comparativo dos métodos geoestatísticos de estimativa e simulação Gaussiana na predição da emissão de CO<sub>2</sub> do solo. Anais. **II Simpósio de Geoestatística em Ciências Agrárias**. p. 1-6, 2011 .ISSN: 2236-2118.
- [12] VALERIANO, M. M. Curvatura vertical de vertentes em microbacias pela análise de modelos digitais de elevação. **Revista Brasileira de Engenharia Agrícola e ambiental**, v. 7, n. 3, p. 539-546, 2003.

## APÊNDICE A - Rotinas R

```
#http://www.leg.ufpr.br/doku.php/disciplinas:geoesalq
# Versão R 2.15.1
# Programa para Análise Geoestatística
# Biase, Adriele Giaretta
# Departamento de Ciências Exatas
# Universidade de São Paulo – ESALQ
# Piracicaba, Brasil
# adrielegbiase@usp.br
# Revisado em: 30-11-2012

## carregando alguns pacotes necessários para a análise geoestatística
require(sp); require(geoR); require(lattice)
require(maptools); require(MASS); require(rgdal)

#-----#
#   VARIÁVEL: precipitação (4 COLUNA)   #
#-----#

setwd("E:/Geoestatística/MG") # redirecionando o diretório
dir()

## carregando a borda do estado de Minas Gerais (MG)
mgbutm <- read.table("MG-bor-utm.txt", head=T)
mgbutm.t <- mgbutm
mgbutm.t$X1 <- mgbutm.t$X1+200000 # Transladando a área p/ trabalhar dentro
head(mgbutm); head(mgbutm.t)      # de um só fuso (coordenadas positivas)

# reduzindo as coodenadas, agiliza o tempo de estimação dos parâmetros no R
mgbutm.t2 <- mgbutm.t
mgbutm.t2$X1 <- (mgbutm.t$X1)/10000 # A escala agora será 1:10000m
mgbutm.t3 <- mgbutm.t2
mgbutm.t3$X2 <- (mgbutm.t2$X2)/10000 #A escala agora será 1:10000m
head(mgbutm.t3); head(mgbutm.t)
ls()

file.show("MG-dat-utm.txt") # todo o conjunto em unidades de utm
dados.o <- read.table("MG-dat-utm.txt", sep=";", head=T)
dados.t <- dados.o
dados.t$X <- dados.t$X+200000 # Transladando p/ ficar dentro de um so fuso
head(dados.o); head(dados.t); length(dados.t$CHUVA)

boxplot(dados.t$CHUVA) # pontos discrepantes
hist(dados.t$CHUVA)
shapiro.test(dados.t$CHUVA)

dados.t[which(dados.t$CHUVA < 600 | dados.t$CHUVA > 2100),] ## outliers!
dados.t2 <- dados.t[which(dados.t$CHUVA > 600 & dados.t$CHUVA < 2100),]
length(dados.t2$CHUVA) # removeu-se os outilier

## visualização dos dados
```

```

boxplot(dados.t2$CHUVA,
  ylab = "Precipitação média anual (mm) em MG (1980–2010)" ## sem outliers!
hist(dados.t2$CHUVA, xlab = "Precipitação", ylab = "Frequência (mm)")
shapiro.test(dados.t2$CHUVA)

#Dividindo LONG e LAT pela mesma cst=10000 para agilizar o
dados.t3<- dados.t2          #tempo de estimação dos parâmetros
dados.t3$X <- (dados.t2$X)/10000
dados.t4<- dados.t3
dados.t4$Y <- (dados.t3$Y)/10000
head(dados.t4);head(dados.t2)

```

```

# Banco Geoestatístico
dG <- as.geodata(dados.t4, coords.col=c(1,2), data.col=4, covar.col=6 )
dG$borders <- mgbutm.t3
plot(dG)

## Análise descritiva dos dados
resumo <- summary(dG); resumo$coords.summary
maior <- resumo$distances.summary[2]; maior
dist1 <- maior*.6; dist1 # distancia recomendada
resumo$covariate.summary

#Estudo gráfico residual levando em consideração tendências
plot(dG, low=T)
plot(dG, low=T, trend=~I(coords[,2]))      # y
plot(dG, low=T, trend=~ALT)                # z
plot(dG, low=T, trend=~I(coords[,2])+ALT)  # y + z
plot(dG, low=T, trend=="1st")              # x + y
plot(dG, low=T, trend=~coords+ALT)         # x + y + z
# x + y + x^2 + y^2
plot(dG, low=T, trend=~coords+I(coords[,1]^2)+I(coords[,2]^2))
plot(dG, low=T, trend=="2nd")              # x + y + x^2 + y^2 + xy
# x + y + x^2 + y^2 + z
plot(dG, low=T, trend=~coords+I(coords[,1]^2)+I(coords[,2]^2)+ALT)

```

```

# Relação entre precipitação e as coordenada geográficas.
x <- dG$coords[,1]
y <- dG$coords[,2]
plot(x,dG$data, xlab="longitude: x*10.000",
      ylab = "Precipitação (mm)" )
x11()
plot(y,dG$data, xlab="latitude: y*10.000",
      ylab = "Precipitação (mm)" )

#Gráfico de círculos, dividido nos quartis
points(dG)
points(dG, pt.div="quarti", cex.min=1, cex.max=1,
        xlab="x*10.000", ylab="y*10.000")
legend(8,770,c("Min-Q1", "Q1-Q2", "Q2-Q3", "Q3-Máx"),
      fill=c("blue", "green", "yellow", "red"))

```

```

# identificar algum outliers spaciais
with( parent.frame(dG$coords), identify(dG$coords[,1],dG$coords[,2]))
# 156 73 328
dG$data[1] # Abadia dos Dourados
dG$data[156] # Francisco SA
dG$data[328] # São João da Vereda
dG$data[101] # Delfim Moreira
dG$data[23] # Bairro Santa Cruz

```

```

# Semi-variograma
resumo <- summary(dG); resumo$coords.summary
resumo$covariate.summary
maior <- resumo$distances.summary[2]; maior
dist1 <- maior*.6; dist1 # sugestão da distância

dist2 <- 50 # distância escolhida
plot(v1 <- variog(dG, max.dist=dist2, trend="1st"),
      xlab="(distância)*10.000")
plot(v2 <- variog(dG, max.dist=dist2, trend=~coords + ALT),
      xlab="(distância)*10.000")
plot(v3 <- variog(dG, max.dist=dist2, trend=~coords + I(coords[,1]^2) +
                  I(coords[,2]^2)), xlab="(distância)*10.000")
plot(v4 <- variog(dG, max.dist=dist2, trend="2nd"),
      xlab="(distância)*10.000")
plot(v5 <- variog(dG, max.dist=dist2, trend=~coords + I(coords[,1]^2) +
                  I(coords[,2]^2) + ALT), xlab="(distância)*10.000")

# Envelopes empíricos
par(mfrow=c(2,3))
v1.mc <- variog.mc.env(dG, obj.variog=v1)
plot(v1, env=v1.mc, main="(a)", xlab="(distância)*10.000")
v2.mc <- variog.mc.env(dG, obj.variog=v2)
plot(v2, env=v2.mc, main="(b)", xlab="(distância)*10.000")
v3.mc <- variog.mc.env(dG, obj.variog=v3)
plot(v3, env=v3.mc, main="(c)", xlab="(distância)*10.000")
v4.mc <- variog.mc.env(dG, obj.variog=v4)
plot(v4, env=v4.mc, main="(d)", xlab="(distância)*10.000")
v5.mc <- variog.mc.env(dG, obj.variog=v5)
plot(v5, env=v5.mc, main="(e)", xlab="(distância)*10.000")
# Existe alguns indicativos de dependência espacial,
# confirmar com o teste posteriormente

```

```

# Foi realizada uma análise preliminar e verificou que o
# modelo spherical é que possui os melhores ajustes (AIC)
# Fixando o modelo spherical no eyefit
ef1 <- eyefit(v4); ef4 # spherical (9758,25), nugget=25370
ef2 <- eyefit(v5); ef5 # spherical (9367,22), nugget=24354
ef3 <- eyefit(v6); ef6 # spherical (8206,18), nugget=25305
ef4 <- eyefit(v7); ef7 # spherical (8943,15), nugget=24146
ef5 <- eyefit(v8); ef8 # spherical (8725,15), nugget=23560

# variofit

```

```

vf1 <- variofit(v1, ini=c(9758,25), nugget=25370,cov.model="spherical"); vf1
vf2 <- variofit(v2, ini=c(9367,22), nugget=24354,cov.model="spherical"); vf2
vf3 <- variofit(v3, ini=c(8206,18), nugget=25305,cov.model="spherical"); vf3
vf4 <- variofit(v4, ini=c(8943,15), nugget=24146,cov.model="spherical"); vf4
vf5 <- variofit(v5, ini=c(8725,15), nugget=23560,cov.model="spherical"); vf5

```

```

# Os respectivos resultados do variofit na ordem:
[modelo] "c(sigma, tau), nugget"
[1] "c(9868.9087288884, 20.412685588397), nug=25177.4391987498"
[2] "c(8104.40652435203, 22.6593375767967), nug=25526.4747233395"
[3] "c(7861.90120131805, 12.0138184241389), nug=24245.4036887095"
[4] "c(7003.49223117441, 11.6414097258205), nug=24522.290142747"
[5] "c(8346.98143879706, 0), nug=22539.551662372"

```

```

# Usando o likfit (chutes iniciais com base no variofit)
args(likfit)

m1 <- likfit(dG, ini=c(9868.9087288884, 20.412685588397),
nug=25177.4391987498 , cov.model="spherical", trend="1st")

m2 <- likfit(dG, ini=c(8104.40652435203, 22.6593375767967),
nug=25526.4747233395 , cov.model="spherical", trend=~coords+ALT)

m3 <- likfit(dG, ini=c(7861.90120131805, 12.0138184241389),
nug=24245.4036887095 , cov.model="spherical", trend=~coords
+ I(coords[,1]^2) + I(coords[,2]^2))

m4 <- likfit(dG, ini=c(7003.49223117441, 11.6414097258205),
nug=24522.290142747 , cov.model="spherical", trend="2nd")

m5 <- likfit(dG, ini=c(8346.98143879706, 0),
nug=22539.551662372, cov.model="spherical", trend=~coords +
I(coords[,1]^2) + I(coords[,2]^2) + ALT)

summary(m1)
summary(m2)
summary(m3)
summary(m4)
summary(m5)

#Análise de resíduos (teste de normalidade) Shapiro–Wilk
names(m1)
residuo1=m1$model.components$residuals
residuo2=m2$model.components$residuals
residuo3=m3$model.components$residuals
residuo4=m4$model.components$residuals
residuo5=m5$model.components$residuals
shapiro.test(residuo1)
shapiro.test(residuo2)
shapiro.test(residuo3)
shapiro.test(residuo4)
shapiro.test(residuo5)

```

```
# Box-Plot para caso de transformações (não foi necessário)
boxcox(dG)
lambda <- boxcox(dG)$x[which(boxcox(dG)$y==max(boxcox(dG)$y))]
lambda
```

```
# Teste da Razão de Verossimilhança (modelos encaixados)
# H0: Modelo 1 mais adequado que o modelo 2
est1=2*(m2$loglik - m1$loglik); est1 # estatística
pchisq(est1, m2$npars-m1$npars, lower.tail=F) #quiquadrado (tabelado)

# H0: Modelo 1 mais adequado que o modelo 3
est2=2*(m3$loglik - m1$loglik); est2 # estatística
pchisq(est2, m3$npars-m1$npars, lower.tail=F) #quiquadrado (tabelado)

# H0: Modelo 1 mais adequado que o modelo 4
est3=2*(m4$loglik - m1$loglik); est3 # estatística
pchisq(est4, m6$npars-m1$npars, lower.tail=F) #quiquadrado (tabelado)

# H0: Modelo 3 mais adequado que o modelo 4
est4=2*(m4$loglik - m3$loglik); est4 # estatística
pchisq(est4, m4$npars-m3$npars, lower.tail=F)#quiquadrado (tabelado)
```

```
# Estudo da Anisotropia
plot(variog4(dG, max.dist=dist2, trend="1st"), legend=T)

plot(variog4(dG, max.dist=dist2, trend=~coords+ALT),
      legend=T, xlab="distância *10.000")

plot(variog4(dG, max.dist=dist2, trend=~coords+I(coords[,1]^2) +
            I(coords[,2]^2)), legend=T)

plot(variog4(dG, max.dist=dist2, trend="2nd"), legend=T)

plot(variog4(dG, max.dist=dist2, trend=~coords+I(coords[,1]^2) +
            I(coords[,2]^2)+ALT), legend=T)
```

```
# definindo um "grid" de predição
grid <- expand.grid(seq(6,125, 1), seq(748,845, 1))
plot(grid, pch=19, cex=0.25, col=4, xlab="longitude (x*10.000)",
      ylab="latitude (y*10.000)")

points(dG, add=T)
points(dG)

#Krigagem com o modelo m2 selecionado
kc1 <- krige.conv(dG, loc=grid, krige=krige.control(obj.m=m2))
image(kc1, col=terrain.colors(21), x.legend=c(67,124),y.legend=c(744,751),
      xlab="longitude (x*10.000)",
      ylab="latitude (y*10.000) ")
```

```
# Validação Cruzada
#? xvalid
VC2 <- xvalid(dG, model=m2)
par(mfrow=c(1,3))
plot(VC2)
lmVC2 <- lm(VC2$data~VC2$predicted); AIC(lmVC2); cor(VC2$data,VC2$predicted)
```