

Comunicado 60

Técnico

Junho, 2004
Campinas, SP

ISSN 1677-8464



Validação de Modelos Geoestatísticos Usando Teste de Filliben: Aplicação em Agroclimatologia

José Ruy Porto de Carvalho¹
Sidney Rosa Vieira²

Tanto o setor público como a administração privada têm exigido informações meteorológicas e climáticas cada vez mais precisas e confiáveis, a fim de estabelecer metas, fazer planejamento estratégico e minimizar custos. Com os investimentos que têm sido feitos por parte do governo federal, a qualidade das previsões de tempo tem melhorado substancialmente ao longo dos últimos cinco anos, auxiliando em muitas atividades econômicas que são sensíveis ao clima. A construção civil, a aviação, a agricultura, assim como a medicina, são influenciados pelo tempo e, conseqüentemente, pelas condições meteorológicas. Especificamente na agricultura, a necessidade das informações meteorológicas é extremamente importante. A duração da estação de crescimento depende do tipo de planta, da temperatura do ar, da disponibilidade de água e da localização geográfica onde a planta está sendo cultivada. A freqüência e a quantidade das precipitações pluviais, juntamente com a capacidade de armazenamento de água no solo, definem a eficiência do uso da água (Bergamaschi, 1992; Vieira & Carvalho, 2001).

O objetivo deste trabalho é validar modelos geoestatísticos aplicados a agroclimatologia através do teste de normalidade de Filliben em resíduos ortonormais com dados de precipitação média anual para o Estado de São Paulo.

Foram usadas quarenta e nove observações de precipitação pluvial média anual obtidas de estações climatológicas

(Fig. 1) abrangendo todo Estado de São Paulo, representando uma área de aproximadamente 248.808,8km² (2,91% do território nacional), no período de 1957 a 1997.



Fig. 1. Localização das estações climatológicas.

O semivariograma é uma função matemática definida para representar o nível de dependência entre duas variáveis aleatórias regionalizadas locais. Ele é usado para modelar valores correlacionados no espaço ou no tempo e é pré-requisito para os métodos de interpolação de krigagem. A modelagem do semivariograma é uma modelagem de cada estrutura de correlação espacial. Considere duas variáveis regionalizadas, X e Y , onde $X = Z(x)$ e $Y = Z(x+h)$.

¹ Estatístico, Pesquisador da Embrapa Informática Agropecuária, Caixa postal 6041, Barão Geraldo - 13083-970 - Campinas, SP. (e-mail: jrui@cnptia.embrapa.br)

² Engenheiro Agrônomo, Pesquisador do IAC - Centro de Pesquisa e Desenvolvimento de Solos e Recursos Ambientais, Caixa Postal 28 - 13001-970 - Campinas, SP. (e-mail: sidney@iac.sp.gov.br)

Neste caso, referem-se ao mesmo atributo por exemplo, o teor x denota uma posição em duas dimensões, com componentes (x_i, y_i) , e h um vetor distância (módulo e direção) que separa os pontos. O nível de dependência espacial entre essas duas variáveis regionalizadas, X e Y , é representado pelo semivariograma, estimado pela seguinte equação:

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

onde $N(h)$ é o número de pares de valores medidos $Z(x)$, $Z(x+h)$, separado pela distância h , se a variável for escalar. O gráfico de $g^*(h)$ versus os valores correspondentes de h , chamado semivariograma, é uma função do vetor h , e, portanto depende de ambos, magnitude e direção de h . Diversos tipos de modelos estão disponíveis na literatura, entretanto os mais usados são os modelos esférico, exponencial e gaussiano que foram ajustados aos semivariogramas, os quais permitem visualizar a natureza da variação espacial das variáveis estudadas, além de serem necessários para outras aplicações, como por exemplo, krigagem.

Ajustar modelos matemáticos aos semivariogramas é um procedimento subjetivo. A qualidade de ajuste pode ser verificada através da técnica de "jack-knifing". Como para cada local têm-se um valor medido e pode-se estimar um outro valor através da krigagem ou co-krigagem, então pode-se calcular a regressão linear entre estes pares de dados e calcular a interseção (a), o coeficiente angular (b), a correlação entre os pares (r) e o erro reduzido com sua média e variância (Vieira et al., 1983; Vieira, 1997). O melhor ajuste se obtém quando os valores obtidos se aproximam dos seguintes valores ideais: $a = 0$, $b = 1$, $r = 1$, média do erro reduzido = 0 e variância do erro reduzido = 1. Definido o modelo matemático, os interpoladores de krigagem são usados e os pontos são estimados com as propriedades de serem não-viciados e com variância mínima (Vieira et al., 1983; Carvalho et al., 2002; Carvalho & Assad, 2003) e ideais para a construção de mapas de isolinhas ou tridimensionais para verificação e interpretação da variabilidade espacial. Entretanto, as condições ideais necessárias pela técnica de "jack-knifing" para definir o modelo que mais se ajusta aos dados são difíceis de se encontrar na prática. O teste de Filliben para resíduos ortonormais pode ser de grande ajuda na determinação de qual o melhor modelo geoestatístico.

Os resíduos ortonormais (Lee, 1994) para p variáveis são definidos como:

$$e_k = \frac{z(x_k) - z_e(x_k)}{S_k} \quad (2)$$

onde e_k ($k = p + 1, \dots, n$) são os resíduos normais, $z(x_k)$ são os valores originais, $z_e(x_k)$ são os valores estimados pelo interpolador de krigagem ordinária (Carvalho & Assad, 2003) e S_k é o desvio padrão da estimativa da krigagem. Os resíduos da equação 2 são chamados resíduos ortonormais, isto é, são não correlacionados e tem variância unitária.

O teste de Filliben (Filliben, 1975) calcula a correlação (r) entre os dados ordenados e a estatística mediana ordenada da distribuição normal com média zero e variância um. Quanto mais próximo r é de 1, mais "normais" são os dados. Se os valores tabelados R , para um determinado tamanho de amostra é maior do que r , a hipótese de que os dados seguem uma distribuição normal é rejeitada para um nível de probabilidade prefixado, α . Praticamente o que se testa é se os resíduos ortonormais seguem uma distribuição normal a um determinado nível de significância.

Para a obtenção do teste, alguns passos são necessários:

1. seja Y o vetor com $n - p$ resíduos ortonormais e X uma amostra ordenada de Y ;
2. calcular a estatística da mediana ordenada m_i , de uma população com distribuição uniforme $[0, 1]$ usando:

$$m_i = 1 - m_n \text{ se } i = 1$$
 ou

$$m_i = (i - 0,3175) / (n + 0,365) \text{ se } i = 2, 3, \dots, n - 1$$
 ou

$$m_i = 0,5 * (1/n) \text{ se } i = n;$$
3. obter a estatística mediana ordenada m_i , de uma distribuição populacional normal $(0, 1)$ usando:

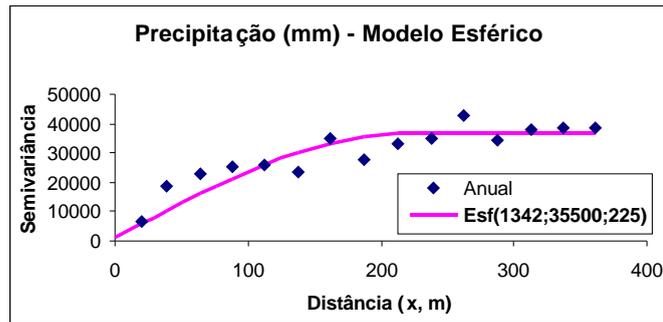
$$m_i = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{m_i} \exp(-t^2/2) dt$$

onde m_i pode ser obtido calculando o algoritmo definido por Abramowitz & Stegun (1964, p. 933), equação 26.2.23.

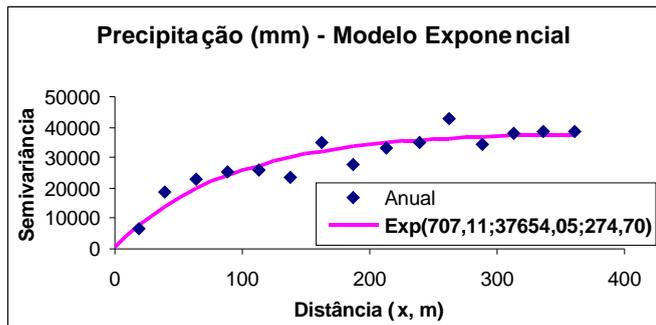
4. calcular $r = \text{corr}(X, M) = \frac{\sum (X_i - X_m)(M_i - M_m)}{\sqrt{[\sum (X_i - X_m)^2] * [\sum (M_i - M_m)^2]}}$

Compare o valor de r com o valor de R tabelado. Se $R > r$, a hipótese de resíduos ortonormais normalmente distribuídos é rejeitada a um nível de probabilidade.

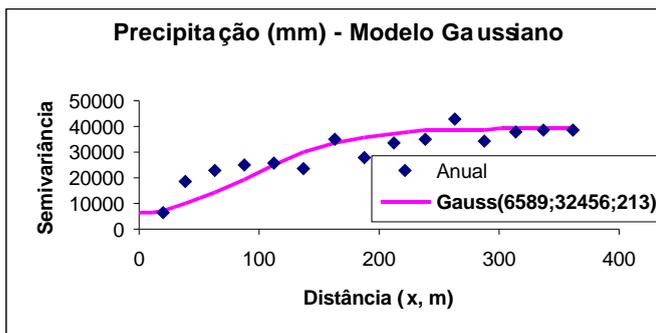
Os dados anuais de precipitação são médias para o período de 1957 a 1997. Todas as análises foram realizadas nestas médias sem se preocupar com as flutuações interanuais. Os semivariogramas experimentais para precipitação anual são apresentados na Fig. 2 e são utilizados para avaliar a dependência espacial da variável em estudo. Os modelos esférico, exponencial e gaussiano foram ajustados aos semivariogramas de precipitação anual média. O exame dos semivariogramas para precipitação anual média, Fig. 2, revela que existe dependência espacial e que apresentam isotropia, ou seja, variabilidade espacial independe da direção escolhida.



A



B



C

Fig. 2. Modelos ajustados aos semivariogramas para precipitação anual média (A = modelo esférico, B = modelo exponencial, C = modelo gaussiano).

Os indicadores obtidos por "Jack-Knifing" (Vieira, 2000) são apresentados na Tabela 1.

Tabela 1. Indicadores de "Jack-Knifing" onde: interseção (a), coeficiente angular (b), correlação entre os pares (r^2) e o erro reduzido com sua média e variância.

Modelo	a	b	r	Erro Média	Reduzido Variância
Esférico	340,5	0,77	0,51	-0,038	1,59
Exponencial	327,2	0,78	0,51	-0,059	1,37
Gaussiano	418,0	0,72	0,50	-0,024	2,86

Pela Tabela 1, o modelo exponencial apresentou valores dos indicadores que mais se aproximaram aos indicadores ideais. Além disso, apresentou um coeficiente de determinação não linear de $R^2 = 0,92$ com a menor Soma de Quadrado de Desvios Ponderados, sendo portanto o modelo escolhido. As estimativas dos parâmetros dos modelos ajustados aos semivariogramas são essenciais na obtenção

dos valores não-amostrados através dos métodos de krigagem. Os valores obtidos através de krigagem são não-variados, têm variância mínima (Vieira, 2000) e são ideais para a construção de mapas de isolinhas para verificação e interpretação da variabilidade espacial.

Os resultados do Teste de Filliben são apresentados na Tabela 2. O valor de R tabelado para 48 resíduos ortonormais é $R = 0,975$ para um nível de probabilidade $\alpha = 5\%$. Como $r > R$, não temos razões para rejeitar o modelo exponencial.

Tabela 2. Teste de Filliben para resíduos ortonormais obtidos com os ajustes dos modelos Esférico, Exponencial e Gaussiano.

Teste de Filliben – Coeficiente de Correlação r		
Esférico	Exponencial	Gaussiano
0,9589	0,9827	0,9632

Conclusões

- Teste de Filliben em resíduos ortonormais demonstrou ser uma poderosa ferramenta para diagnosticar a validade dos modelos geoestatísticos estimados; e
- para o conjunto de dados de precipitação média anual do Estado de São Paulo, o modelo que melhor se ajustou ao semivariograma foi o Exponencial.

Referências Bibliográficas

- ABRAMOWITZ, M.; STEGUN, I. A. Handbook of mathematical functions. Washington, D. C.: National Bureau of Standards, 1964. 1058 p.
- BERGAMASCHI, H. Agrometeorologia aplicada à irrigação. Porto Alegre: UFRGS, 1992. 125 p.
- CARVALHO, J. R. P. de; ASSAD, E. D. Análise espacial da precipitação pluviométrica no Estado de São Paulo: comparação de interpoladores. In: CONGRESSO BRASILEIRO DE ENGENHARIA AGRÍCOLA, 32., 2003, Goiânia. [Anais...]. Goiânia: SBEA, 2003. CD-ROM.
- CARVALHO, J. R. P. de; SILVEIRA, P. M. da; VIEIRA, S. R. Geoestatística na determinação da variabilidade espacial de características químicas do solo sob diferentes preparos. Pesquisa Agropecuária Brasileira, Brasília, DF, v. 37, n. 8, p. 1151-1159, ago. 2002.
- FILLIBEN, J. J. The probability plot correlation coefficient test for normality. Technometrics, v. 17, n. 1, p. 111-117, 1975.
- LEE, S. -I. Validation of geostatistical models using the Filliben test of orthonormal residuals. Journal of Hydrology, v. 158, p. 319-332, 1994.

VIEIRA, S. R. Geoestatística em estudos de variabilidade espacial do solo. In: NOVAIS, R. F. de; ALVAREZ V., V. H.; SCHAEFER, C. E. G. R. (Ed.). Tópicos em ciência do solo. Viçosa: Sociedade Brasileira de Ciência do Solo, 2000. v. 1, p. 1-54.

VIEIRA, S. R. Variabilidade espacial de argila, silte e atributos químicos em uma parcela experimental de um latossolo roxo de Campinas (SP). *Bragantia*, Campinas, v. 56, n. 1, p. 1-17, 1997.

VIEIRA, S. R.; CARVALHO, J. R. P. de. Estudo da periodicidade temporal de chuvas em bacia hidrográfica dos Rios Turvo / Grande - uma proposta. Campinas: Embrapa Informática Agropecuária, 2001. 17 p. (Embrapa Informática Agropecuária. Documentos, 10).

VIEIRA, S. R.; HATFIELD, J. L.; NIELSEN, D. R.; BIGGAR, J. W. Geoestatistical theory and application to variability of some agronomical properties. *Hilgardia*, Berkeley, v. 51, n. 3, p. 1-75, 1983.

Comunicado Técnico, 60

Ministério da Agricultura,
Pecuária e Abastecimento
Governo Federal

Embrapa Informática Agropecuária
Área de Comunicação e Negócios (ACN)
Endereço: Caixa Postal 6041 - Barão Geraldo
13083-970 - Campinas, SP
Fone: (19) 3789-5743
Fax: (19) 3289-9594
e-mail: sac@cnptia.embrapa.com.br

1ª edição on-line - 2004

ÓTodos os direitos reservados.

Comitê de Publicações

Presidente: *Luciana Alvim Santos Romani*
Membros Efetivos: *Carla Geovana Macário, José Ruy Porto de Carvalho, Marcia Izabel Fugisawa Souza, Marcos Lordello Chaim, Suzilei Almeida Carneiro.*
Suplentes: *Carlos Alberto Alves Meira, Eduardo Delgado Assad, Maria Angelica Andrade Leite, Maria Fernanda Moura, Maria Goretti Gurgel Praxedis.*

Expediente

Supervisor editorial: *Ivanilde Dispató*
Normalização bibliográfica: *Marcia Izabel Fugisawa Souza*
Editoração eletrônica: *Área de Comunicação e Negócios*