

TESTE DE FILLIBEN PARA VALIDAR MODELOS GEOESTATÍSTICOS APLICADOS A AGROCLIMATOLOGIA

JOSÉ RUY PORTO DE CARVALHO¹; SIDNEY ROSA VIEIRA²

Escrito para apresentação no
XXXIII Congresso Brasileiro de Engenharia Agrícola
02 a 06 de Agosto de 2004 - São Pedro - SP

RESUMO: O objetivo deste trabalho é validar modelos geoestatísticos aplicados a agroclimatologia através do teste de normalidade de Filliben em resíduos ortonormais, usando quarenta e nove (49) observações de precipitação pluvial média anual obtidas de estações climatológicas abrangendo todo Estado de São Paulo, representando uma área de aproximadamente 248.808,8 km², no período de 1957 a 1997. Os resultados da aplicação do teste de Filliben comprovaram seu poder para validar modelos geoestatísticos sendo que o modelo exponencial foi o que melhor ajustou para os dados em questão.

PALAVRAS-CHAVE: Precipitação pluvial, Resíduos ortonormais, Variabilidade espacial

Filliben test to validate geostatistical models applied to agroclimatology

ABSTRACT: The objective of this work is to validate geostatistical models applied to agroclimatology through the Filliben normality test in orthonormal residuals, using forty nine (49) annual precipitation average observations obtained from climatological stations including all State of São Paulo, representing an area of approximately 248.808,8 km², in the period from 1957 to 1997. The results of the application of the Filliben test proved its power to validate geostatistical models and the exponential model was what best adjusted for the data in question.

KEYWORDS: Pluvial precipitation, Orthonormal residuals, Spatial variability

INTRODUÇÃO: 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 tem sido feito 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 frequência 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.

MATERIAL E MÉTODOS: Foram usadas quarenta e nove (49) observações de precipitação pluvial média anual obtidas de estações climatológicas abrangendo todo Estado de São Paulo, representando uma área de aproximadamente 248.808,8 km² (2,91% do território nacional), no período de 1957 a 1997.

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

1- Estatístico, Pesquisador III, Embrapa Informática Agropecuária, Embrapa, Barão Geraldo - Campinas-SP, 19 - 3789.5797, jruy@cnptia.embrapa.br

2- Engenheiro Agrônomo, Pesquisador, Centro de Solos, IAC, Campinas-SP

semivariograma é uma modelagem de cada estrutura de correlação espacial. Considere duas variáveis regionalizadas, X e Y, onde $X = Z(x)$ e $Y = Z(x+h)$. 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 equação (1), onde $N(h)$ é o número de pares de valores medidos $Z(x_i)$, $Z(x_i+h)$, separado pela distância h, se a variável for escalar. O gráfico de gerado pela equação 1 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 pela equação 2, onde $e_{p(k)}$ ($k=p+1, \dots, n$) são os resíduos normais, $z(x_k)$ são os valores originais, $ze(x_k)$ são os valores estimados pelo interpolador de krigagem ordinária (Carvalho & Assad, 2003) e $\sigma(k)$ 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 pré-fixado nível de probabilidade alfa. 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: $m_i = 1 - m_{n-i}$ se $i=1$ ou $(i-0,3175)/(n+0,365)$ se $i=2,3, \dots, n-1$ ou $0,5^{**}(1/n)$ se $i=n$;
3. Obter a estatística mediana ordenada M_i , de uma distribuição populacional normal (0,1) usando: a equação 3, onde M_i pode ser obtido calculando o algoritmo definido em (Abramowitz et al.; 1964) página 933, equação 26.2.23.
4. Calcular $r = \text{corr}(X, M) =$ equação 4
5. 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 alfa de probabilidade.

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

Equação 1

$$\varepsilon_k = \frac{Z(X_k) - Z_m(X_k)}{\sigma_k} \quad (2)$$

Equação 2

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

Equação 3

$$\frac{\sum (X_i - X_m)(M_i - M_m)}{\sqrt{\sum (X_i - X_m)^2 \sum (M_i - M_m)^2}}$$

Equação 4

RESULTADOS E DISCUSSÃO: 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 Figura 1 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, Figura 1, revela que existe dependência espacial e apresentam isotropia, ou seja, variabilidade espacial independe da direção escolhida.

Os indicadores obtidos por "Jack-Knifing" (Vieira, 2000) são apresentados na Tabela 1. O ajuste ideal é obtido quando os valores se aproximam dos seguintes valores ideais: a = 0, b = 1, r = 1, média do erro reduzido = 0 e variância do erro reduzido = 1.

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 R²=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 viciados, 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 alfa = 5%. Como r > R, não temos razões para rejeitar o modelo exponencial.

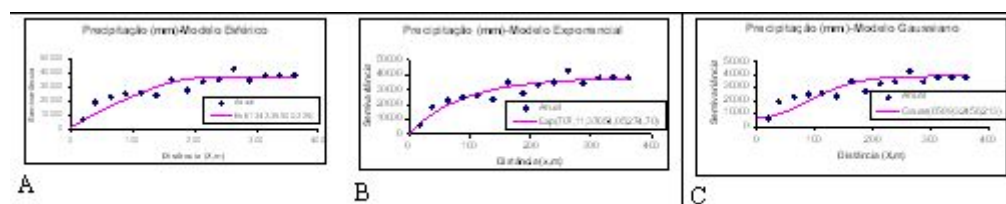


Tabela 1 – Indicadores de "Jack-Knifing" onde: interseção (a), coeficiente angular (b), correlação entre os pares (r²) e o erro reduzido com sua média e variância. Modelos ajustados aos semivariogramas para precipitação anual média (A = modelo esférico, B = modelo exponencial, C = modelo gaussiano)

Modelo	a	b	r	Erro	Reduzido
				Média	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

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: Modelos usados para estimar funções espaciais necessitam de validação antes de serem aplicados;

O Teste de Filliben em resíduos ortonormais demonstrou ser uma poderosa ferramenta para diagnosticar a validade dos modelos geoestatísticos estimados;

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. National Bureau of Standards, Washington, D. C. 1058p., 1964.
- BERGAMASCHI, H. Agrometeorologia aplicada à irrigação. Porto Alegre: UFRGS, 1992. 125p.
- 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, v. 37, n. 8, p. 1151-1159, ago. 2002.
- CARVALHO, J. R. P. de; ASSAD, E. D. Análise espacial da precipitação pluviométrica no Estado de São Paulo; comparação de interpoladores. In: XXXII Congresso Brasileiro de Engenharia Agrícola, 2003, Goiânia, GO. Anais, Goiânia, 2003. CD-ROM.
- LEE, S.-I. Validation of geostatistical models using the Filliben test of orthonormal residuals. Journal of Hydrology, v.158, p.319-332, 1994.
- FILLIBEN, J. J. The probability plot correlation coefficient test for normality. Technometrics, v. 17, n. 1, p. 111-117, 1975.
- 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. Geoestatística em estudos de variabilidade espacial do solo. In: NOVAIS, R. F. de; ALVARZ V., V. H.; SCHAEFER, C. E. G. R. Tópicos em Ciência do Solo, Sociedade Brasileira de Ciência do Solo, v. 1, p. 1-54, 2000.
- 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, 17p : - (Documentos/Embrapa Informática Agropecuária ; 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.