

Análise espacial aplicada na determinação de valor de terrenos em Campo Mourão (Pr)

Dieter Randolf Ludewig, EPA, UNESPAR/Campus de Campo Mourão

dludewig@fecilcam.br

Resumo: O presente estudo objetiva determinar um estimador para o valor dos imóveis urbanos da cidade de Campo Mourão (Pr) com a utilização da tecnologia SIG. Uma vez encontrado como principal fator na determinação do valor de um imóvel urbano, a sua localização bem como a dependência espacial deste com os demais imóveis em escala, usando técnicas de Geoprocessamento e análise espacial, procurou-se estabelecer através de técnicas de interpolação, um zoneamento por isolinhas que permitisse estabelecer um parâmetro de valor dentro de um espectro estocástico. Este trabalho procura estabelecer a correlação do valor do m² dos terrenos e a sua distribuição espacial.

PALAVRAS-CHAVES: Geoestatística; Geoprocessamento; SIG.

1. Introdução

A aplicação de instrumentais de geoestatísticas aliados ao geoprocessamento de imagens proporciona ao tomador de decisões, instrumentos facilitadores destas decisões. A disponibilidade no mercado, de softwares de gestão espacial com a mais variada gama de softwares complementares só necessitam de pessoal técnico e capacitado para produzir informações de alto nível técnico, diferencial que as técnicas convencionais de gestão de informações geograficamente distribuídas, não proporciona. A aplicação de Sistemas de Informação Geográfica (SIG) na aplicação de políticas públicas não é nenhuma novidade, no entanto tem sido pouco utilizado por conferir parâmetros técnicos à tomada de decisão.

Neste estudo, optamos por uma seleção técnica das teorias relacionado a Estatística e Geoestatística bem como a aplicação dos métodos com a utilização do software SURFER 12(Golden Software Inc. 2012) para a geração dos mapas de contorno e de isolinhas, Geocac (Unioeste 2005) para os cálculos dos parâmetros geoestatísticos, e para a conversão das coordenadas geográficas em UTM utilizamos a Calculadora Geográfica produzida pela Unesp – Campus de Presidente Prudente encontrada no site <http://www.dpi.inpe.br/calcula/>

2. Formatação a ser utilizada

2.1 Considerações gerais

2.1.1 - A Teoria das Variáveis Regionalizadas

A Teoria das Variáveis Regionalizadas é o estudo de funções que variam de um lugar para outro no espaço com certa aparência de continuidade. Isto é, são variáveis cujos valores são

relacionados de algum modo com a posição espacial que ocupam. Esta teoria diz que algumas medidas podem ser vistas como uma realização de uma função aleatória ou processo estocástico.

A continuidade atribuída às variáveis regionalizadas se manifestam sobretudo pela tendência de tomar em valores mais próximos em dois pontos amostrais, quando menos afastados estejam os referidos pontos.

A teoria das variáveis regionalizadas desenvolvidas por Matheron leva em consideração que as variáveis são aleatórias, já que os valores numéricos observados podem variar consideravelmente de um ponto a outro do espaço e são espaciais, já que as variáveis não são inteiramente independentes.

A variação espacial de uma variável regionalizada pode ser expressa pela soma de três componentes: uma componente estrutural, associada a um valor médio constante ou a uma tendência constante ($m(s)$); uma componente aleatória, espacialmente correlacionada ($\delta(s)$) e um ruído aleatório ou erro residual ($\varepsilon(s)$).

Se s representa uma posição em uma, duas ou três dimensões, então o valor da variável Z , em s , é dado por $Z(s) = m(s) + \delta(s) + \varepsilon(s)$ onde, pela sua independência espacial pressuposta, pode-se fazer $\varepsilon(s) = \varepsilon_i$.

2.1.2 - Geotecnologias

Segundo MOCELIN (2003), as tecnologias conhecidas como "geoprocessamento", são o conjunto de tecnologias para coleta, processamento, análise e disponibilização de informação com referência geográfica e são compostas por soluções em hardware, software e peopleware que juntos se constituem em poderosas ferramentas para tomada de decisão.

Dentre as geotecnologias estão os GIS - Sistemas de Informação Geográfica, Cartografia Digital, Sensoriamento Remoto por Satélites, Sistema de Posicionamento Global (ex. GPS), Aerofotogrametria, Geodésia e Topografia Clássica, dentre outros.

O geoprocessamento é o conceito mais abrangente das geotecnologias e representa qualquer tipo de processamento de dados georeferenciados, enquanto o GIS processa dados gráficos e não gráficos com ênfase nas análises espaciais e modelagens de superfícies.

Com a utilização da tecnologia disponível no software Surfer 12, geramos o mapa do contorno de Campo Mourão utilizando os dados georeferenciados contidos na Lei Complementar n°. 30/2013 de 29 de novembro de 2013 que proporcionou uma visão apurada do novo contorno da cidade de Campo Mourão contidas na figura 01 a seguir.

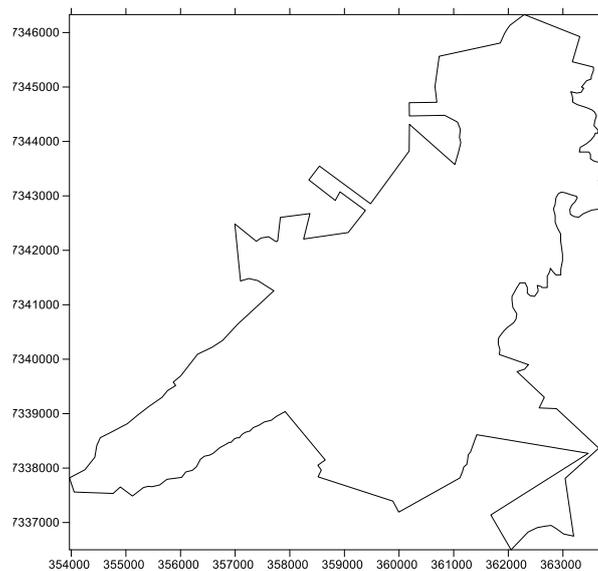


Figura 01 - Mapa do contorno de Campo Mourão

2.1.3 - Estatística

Para GARCIA (1997), a estatística clássica resume-se como sendo aquela que utiliza a média e o desvio-padrão para representar um conjunto de dados, e baseia-se na hipótese principal de que as variações de um local para outro são aleatórias e independentes.

Sokal & Rohlf (1969) citados por GUIMARÃES (1993) afirmam que os valores esperados num dado ponto são estimados levando-se em conta a média da população e a variação dos valores medidos em torno da média. Esta estimativa pressupõe também que as amostras sejam independentes e que possuam distribuição normal. Mostraram ainda que a normalidade dos dados é um pré-requisito básico para o uso da estatística de Fisher, e que pode ser verificado por testes de normalidade como os de Shapiro-Wilk e Anderson-Darling, juntamente com as medidas de forma que são os coeficientes assimetria e curtose.

VIEIRA (1998), acrescenta que se a distribuição de frequência não for normal, deve-se realizar a transformação de dados para que estes tornem-se normais, e o uso da estatística clássica possa a ser utilizado.

Uma etapa importante e que deve ser cumprida, refere-se a análise exploratória dos dados. Os diagramas de ramos e folhas e principalmente os gráficos de box-plot são instrumentos importantes para se detectar dentre os dados os outliers. Outliers ou pontos extremos são dados que estão afastados demais do conjunto de dados. Num gráfico de box-plot, além de podermos observar estes pontos, ele ainda nos indica a mediana, a dispersão dos dados, a assimetria e a distribuição dos dados.

Conforme ressalta VIEIRA (1998), o uso de métodos estatísticos clássicos como análise de variância e coeficiente de variação, só pode ser feito quando não existir correlação de amostras com a distância. Entretanto, segundo GUIMARÃES (1993), pesquisas tem demonstrado que a quantificação de características e propriedades do solo são influenciadas pela variabilidade espacial e, portanto, existe um raio de dependência entre amostras onde a média não é representativa da área de estudo. Neste caso, é mais indicado o uso de estimativas que levem em consideração a posição espacial dos valores amostrados, que é o caso da geoestatística.

2.1.4 - Semivariograma

Se for provada a correlação espacial dos dados experimentais, então a hipótese de independência das mesmas fracassa, desta forma a geoestatística é de grande importância para estudar estes casos.

A semivariância por definição de um processo intrinsecamente estacionário é definida em (8) como: $\gamma(h) = \frac{1}{2} E[(Z(s_i) - Z(s_i+h))^2]$.

Seja $\hat{\gamma}(h)$ a função semivariância estimada de $\gamma(h)$ que pode ser a semivariância de Matheron, (MATHERON, 1963), que tem a forma:

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(s_i) - Z(s_i+h)]^2,$$

onde, $N(h)$ é o número de pares de valores medidos $Z(s_i)$, $Z(s_i+h)$, separados por distância h .

Desta forma, para a obtenção de tal estimador não é necessário o conhecimento da distribuição do processo $Z(s_i)$. Mas este estimador também vem sendo contestado e colocado à prova por diversos autores.

Outros estimadores têm sido propostos buscando uma maior qualidade de estimação. CRESSIE & HAWKINS (1980), comparando 10 estimadores para seis conjuntos de dados simulados para uma transeção segundo um processo autoregressivo, concluem que, sob normalidade, o estimador de Matheron é o mais estável. Porém, tal fato não ocorre quando são simuladas distribuições de causas pesadas. O artigo mostra ainda que a média aritmética da raiz quadrada de $[Z(s_i) - Z(s_i+h)]^2$ fornece uma estimativa da semivariância razoavelmente robusta e

estável, com distribuição próxima da normal. Surge daí um estimador que vem sendo tratado como “O estimador de Cressie & Hawkins” pela literatura (WEBSTER, 1985). CRESSIE e HAWKINS (1980) apresentam o estimador da função semivariância $\gamma(h)$ da forma:

$$\hat{\gamma}(h) = \frac{1}{2} \frac{\left[N(h)^{-1} \sum_{i=1}^{N(h)} \sqrt{|z(x_i + h) - z(x_i)|} \right]^4}{0,457 + \frac{0,497}{N(h)}}$$

Além dos estimadores de semivariância de Matheron e Cressie & Hawkins, a literatura apresenta o estimador relativo de Pairwise (ISAAKS & SRIVASTAVA, 1989; LI & LAKE, 1994).

O estimador relativo de Pairwise é utilizado para distribuições assimétricas e estima a semivariância através de uma média quadrada, de dois valores amostrados, como uma média local, dado por:

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

A principal diferença entre o estimador relativo de Pairwise e o estimador de Matheron é o denominador, que serve para reduzir a influência de pontos discrepantes. O único inconveniente deste estimador é que o denominador deve ser diferente de zero e quando a média se aproxima de zero, poderá ocorrer semivariogramas com ‘saltos’ nas semivariâncias (ISAAKS & SRIVASTAVA, 1989).

CRESSIE (1993) especifica que, no caso do semivariograma, se afastar do modelo significa uma pequena contaminação de um processo gaussiano.

Para a construção de um semivariograma experimental, além da escolha de um estimador para a semivariância, outras escolhas precisam ser feitas. Na prática o semivariograma experimental se calculará até 1/4, 1/3, 1/2 do campo geométrico, depois deste valor o semivariograma perde o significado segundo GUERRA, (1988) e CLARK, (1979). É preciso também escolher o valor h para os quais as semivariâncias experimentais serão estimadas (*lags*), e através de quais pontos amostrais as semivariâncias serão estimadas. Se a amostragem é regular, uma escolha natural para os *lags* é a própria distância entre pontos amostrais, Nesse caso cada semivariância experimental pode ser calculada com base em todos os pares de pontos amostrais separados por um mesmo vetor h . No caso de amostragens irregulares, é possível que haja um número muito pequeno de pares de pontos amostrais separados por exatamente o mesmo vetor h . Nesse caso, McBRATNEY & WEBSTER (1986) recomendam que seja admitida uma tolerância Δh na distância e uma tolerância $\Delta\theta$ na direção, de forma a se atender ao requisito de um número mínimo de pontos amostrais para a estimativa da semivariância a cada *lag*.

O número de pares com que será calculado o semivariograma experimental desempenha um papel importante, sendo necessário segundo WOLLENHAUPT et al. (1997), JOURNAL & HUIJBREGTS (1978) e GUERRA (1988), pelo menos um mínimo de 30 pares, enquanto WEBSTER & OLIVER (1990) recomendam um número mínimo de 100 pares.

Um semivariograma experimental com características muito próximo do ideal. O seu padrão representa o que, intuitivamente, se espera de dados de campo, isto é, que as diferenças $[Z(s_i) - Z(s_i + h)]$ decresçam a medida que h , a distância que os separa decresce. É esperado que observações mais próximas geograficamente tenham um comportamento mais semelhante entre si do que aquelas separadas por maiores distâncias. Desta maneira, é esperado que $\gamma(h)$ aumente com a distância h até atingir um valor máximo no qual se estabiliza. Os parâmetros do seminariograma, que são:

Alcance (a): distância dentro das quais as amostras apresentam-se correlacionada espacialmente;

Patamar (C): é o valor da qual $\gamma(h)$ se estabiliza e a distância h corresponde é a alcance (a). Deste ponto em diante, considera-se que não existe mais dependência espacial entre as amostras, porque a variância da diferença entre pares de amostras ($VAR[Z(s_i) - Z(s_i + h)]$) torna-se invariante com a distância;

Efeito Pepita (C_0): idealmente $\gamma(0) = 0$. Entretanto, na prática, à medida que h tende para zero, $\gamma(h)$ se aproxima de um valor positivo chamado Efeito Pepita (“*nugget effect*”), que revela a descontinuidade do semivariograma para distâncias menores do que a menor distância entre as amostras. Parte desta descontinuidade pode ser também devida a erros de medição (ISAAKS & SRIVASTARA, 1989), mas é impossível quantificar se a maior contribuição provém dos erros de medição ou da variabilidade de pequena escala não captada pela amostragem. A interpretação dada para o efeito pepita é de erros de medição ou de variabilidade em pequena escala (CRESSIE, 1993);

Contribuição (C_1): é a diferença entre o patamar (C) e o Efeito pepita (C_0). O parâmetro C_1 é conhecido como *Sill*.

O patamar (C), o alcance (a) e o efeito pepita (C_0) são os parâmetros por meio dos quais se busca quantificar a dependência espacial da propriedade em estudo.

Se o semivariograma, ao invés de ser crescente e dependente de h como mostra a Figura 3, for constante e igual ao patamar para qualquer valor h , então se tem um efeito pepita puro ou ausência total de dependência espacial. Isto significa que o alcance a , para dados em questão, é menor do que o menor espaçamento entre amostras. Para estes dados, tem-se uma distribuição espacial completamente aleatória, e a única estatística aplicável, é a Estatística Clássica.

Com a utilização do referencial teórico acima, realizamos o desenvolvimento de um modelo experimental de avaliação de preços de mercado de imóveis com a utilização de técnicas de georeferenciamento de dados (tecnologia GIS) na cidade de Campo Mourão que apresentamos no anexo 02. A geração do mapa de isolinhas apresentado na figura 2 foi gerado utilizando-se dados tratados estatisticamente como segue:

- a) Intervalo de interpolação dos dados georeferenciados: Leste – mínimo 353967 e máximo 363803. Norte – mínimo 7336499 e máximo 7346330.
- b) Variância da amplitude dos valores - 50424,1762238 para determinar o efeito pepita.
- c) Contribuição (C_1): Diferença entre patamar – 80120
- d) Alcance (sill): 6213
- e) Estatística de dados regionalizados: O estimador de Cressie & Hawkins” pelo método exponencial

3. Considerações Finais

Com a aplicação dos dados levantados e estatisticamente tratados, geramos o mapa constante da Figura 02, onde podemos ver claramente a dependência espacial dos valores dos terrenos na cidade de Campo Mourão bem como a sua regionalização.

Os mapas de isolinhas são instrumentos de visualização dos dados geograficamente distribuídos e são instrumentos poderosos para o poder público gerar políticas públicas visando o desenvolvimento urbano com base em parâmetros de mercado.

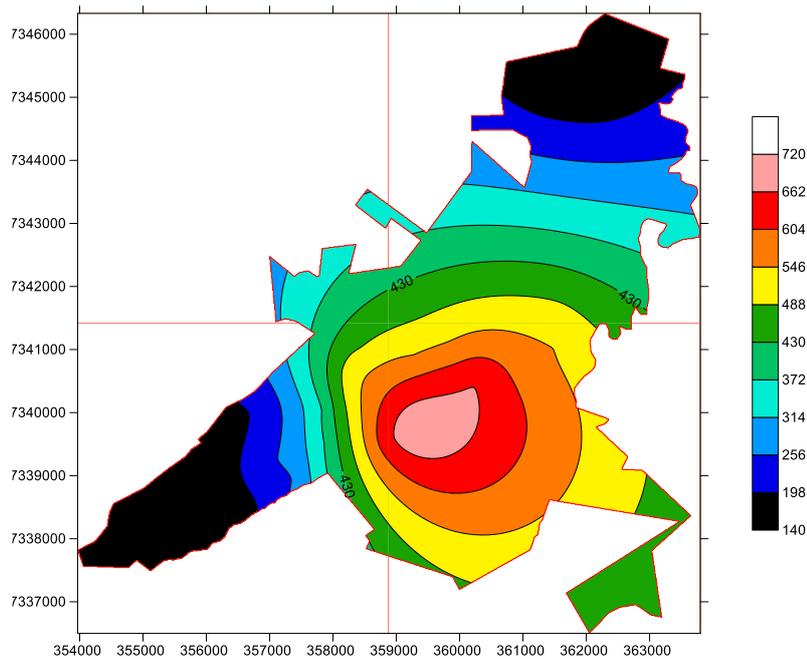


Figura 02 – Mapa do modelo experimental de avaliação do valor dos imóveis (terrenos) da cidade de Campo Mourão

Referências

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Referências bibliográficas**. NBR 6023. Rio de Janeiro, 2002. 22p.

GOLDEN SOFTWARE, INC. **Surfer for Windons-User guide**. Colorado: Golden Software, Inc., 2012. 340p.

ISSAKS, E. H.; SRIVASTAVA, R. M. **Applied Geostatistics**. New York: Oxford University Press. 1989. 561p.

JOHNSON, R.A.; WICHERN, D.W. **Applied multivariate statistical analysis**. New Jersey: Prentice-Hall, 1982. 593p.

MUCELIN, C. A. **Estatística Elementar e Experimental Aplicada às Tecnologias**. Medianeira-Pr: 2003. 245p.

SCHIMALSKI M. B.; LOCH C.; SIEVERS R. **Metodologia para a Obtenção de Mapas de Produtividade**. COBRAC 2004 – Congresso Brasileiro de Cadastro Técnico Multifinalitário UFSC Florianópolis, 10 a 14 outubro de 2004.

SILVA, E. A. A.; OPAZO, M. A. U.; VILAS BOAS, M. A.; LAMPARELLI, R. **GEOCAC** - Software Geoestatístico desenvolvido pelo LEA – Laboratório de Estatística Aplicada da UNIOESTE – Cascavel, última versão atualizada em 04-07-2007.

UNGARO, M. R. G.; PECHE FILHO, A.; LINO, A. C. L.; STORINO, M. **Uso de SIG e de Mapas Temáticos de Estande e Produção de Grãos na Avaliação de Lavoura de Girassol (*Helianthus annuus L.*)** Eng. Agríc., Jaboticabal, v.18, n.3, p.73-79. mar. 1999.

