

UM MODELO ESPACIAL BASEADO EM MÉTODO DE CLASSIFICAÇÃO FUZZY, APLICADO AO MAPEAMENTO DE RISCO À EROSIÃO.

Marcos César Ferreira

¹Instituto de Geociências, UNICAMP, Brasil
macferre@ige.unicamp.br

Daniilo Trovó Garófalo

²Doutorando, Instituto de Geociências, UNICAMP, Brasil

Cassiano Messias

²Doutorando, Instituto de Geociências, UNICAMP, Brasil

Marta Marujo Ferreira

³Instituto de Ciências da Natureza, UNIFAL-MG, Brasil

ABSTRACT

The aim of this paper is to present a spatial decision model based on logistic regression, fuzzy classification and GIS-based techniques, applied to erosion risk mapping. The spatial database was composed by following maps: NDVI vegetation index (1); terrain slope calculated from ASTER-GDEM2 (2); structural lineaments density (3); road density (4) and a map of 729 gullies sites surveyed in the study area (5), obtained by Google images visual analysis and field trip. The values of maps 1 to 4 were classified in 10 classes using the quantil method. Total of gullies occurring in each class was calculated using overlay between one each of the maps 1 to 4, and the map 5. Odd ratios (OR) values indicating the chances in favor of a gully event in relation to the chances against it, in each class map, were estimated using logistic regression. Then, OR curves were converted in fuzzy values using membership functions (Mf). The erosion risk in each pixel was estimated using the weighted sum of the following fuzzified variables: vegetation index, terrain slope, road density and structural lineaments density.

Keywords: Fuzzy classification, GIS, risk map, gullies, logistic regression

Palavras Chave: classificação fuzzy, SIG, mapa de riscos, ravinas, regressão logística

Introdução

A maioria dos algoritmos de classificação de dados disponíveis em sistemas de informação geográfica (SIG) utilizam regras de decisão booleanas para definição dos limites das classes do mapa. Estas regras convencionais baseiam-se apenas em duas possibilidades de classificação dos dados mapeados: pertencer ou não pertencer a uma determinada classe do mapa. Contudo, a maioria das variáveis ambientais perde muito do seu detalhamento e de sua complexidade natural quando mapeada de acordo com a lógica de decisão booleana. Este problema se amplifica em processos de modelagem ambiental, em que, dados de mapas temáticos e imagens orbitais são combinados com objetivo de dar suporte a um processo decisório de atribuição de áreas a uma determinada classe de de risco.

Uma das alternativas para minimização destes problemas é a utilização de um método de classificação contínua. Nos processos de decisão contínua, cada valor de uma variável ambiental tem maior afinidade a pertencer a uma classe A e menor afinidade a pertencer a uma classe B ou C do mesmo mapa, e, assim por diante, para todas as n classes do deste mapa. No processo classificatório fuzzy a afinidade de um valor a uma classe específica é estimada por meio de uma função de afinidade em relação a uma determinada classe.

A função de afinidade $f_A(x)$ associa um número real entre 0,0 e 1,0 a cada pixel x do mapa e indica a possibilidade de este x pertencer a classe A. A lógica utilizada neste tipo método de

classificação é conhecida como lógica fuzzy (Zadeh, 1965; Burrough, 1989; Burrough et al., 1992; Wang and Hall, 1996). Na lógica fuzzy os limites espaciais entre duas ou mais classes de um mapa são representados por zonas transicionais. Os valores de uma variável são distribuídos continuamente no mapa e representados por uma legenda com valores reais que variam entre 0,0 e 1,0.

Outro desafio comum à modelagem de riscos é a determinação dos pesos a serem atribuídos às variáveis ambientais envolvidas no modelo. De maneira geral, os pesos são escolhidos por meio de processos hierárquicos subjetivos baseados no senso comum, ou através de métodos de regressão linear baseados na relação entre uma variável de resposta X e um regressor Y. No entanto, na modelagem de risco, o regressor é uma variável binária ou dicotômica (ocorrência, Y = 1; não-ocorrência Y = 0). Neste caso, onde os valores de uma das variáveis são presença ou ausência de um fenômeno, o método de regressão logística é mais apropriado que o método de regressão linear (Panik, 2009). O objetivo geral deste trabalho é propor um modelo de decisão espacial baseado em regressão logística, lógica fuzzy e técnicas de SIG, aplicado ao mapeamento de áreas de risco à erosão em áreas montanhosas tropicais. Como objetivo específico, propomos a desenvolver um procedimento para atribuição de pesos a variáveis ambientais utilizadas em modelos de risco. Esta pesquisa foi desenvolvida na área de proteção ambiental (APA) Fernão Dias, localizada na Serra da Mantiqueira, sul do estado de Minas Gerais, Brasil (Figura 1).



Figure 1 - Localização da Área de Proteção Ambiental Fernão Dias, no sul do estado de Minas Gerais, Brasil (Fonte: Garófalo et al., 2012).

Material e Métodos

Uma base de dados composta por quatro mapas raster de variáveis ambientais que influenciam a erosão em áreas tropicais foi utilizada para desenvolvimento deste modelo. Estes mapas são os seguintes: índice de vegetação NDVI (VI); declividade do terreno (SL); densidade de lineamentos estruturais (LD) e densidade de estradas (RD). Os dados utilizados para o mapa VI foram gerados a partir de imagens Landsat TM5 de 21 de abril de 2011, nas bandas espectrais do vermelho e infra vermelho próximo. A imagem do índice de vegetação foi obtida pelo algoritmo NDVI, disponível no ENVI 4.8, e, posteriormente exportada para o SIG ArcGIS 10 (ESRI, 2010) no formato raster. O mapa SL foi produzido a partir de dados altimétricos do sensor ASTER GDEM2, processados no módulo de análise espacial do ArcGIS 10. Os mapas LD e RD foram produzidos a partir de interpretação de cartas topográficas 1:50.000 que cobrem a área de

estudo. Em seguida, estes foram exportados como arquivos vetoriais para o ArcGIS 10 e transformados, por meio do algoritmo Kernel, em mapas de densidade.

Os valores dos pixels dos mapas VI, SL, LD e RD foram agrupados em 10 classes por meio do método de classificação quantil. Este procedimento foi utilizado para minimizar o efeito das diferenças entre os tamanhos das áreas das classes dos mapas. Além dos mapas das quatro variáveis ambientais já citadas, um quinto mapa, contendo 729 pontos correspondentes ao centróide das linhas de ravinas (Garófalo et al., 2012), foi adicionado a base de dados. Este mapa foi obtido por meio de análise visual de imagens de alta resolução da plataforma Google Earth.

A relação entre a ocorrência de ravinas e os valores das variáveis ambientais da base de dados foi estimada por meio de operações de overlay entre o mapa de pontos das ravinas e cada um dos mapas VI, LD, SD e RD classificados. Assim, foi calculado o número de ravinas que ocorreram em cada uma das 10 classes dos mapas das variáveis ambientais. Em seguida foi calculado os valores do odd ratio (OR), que indicam a chance em favor da ocorrência de um evento de ravina (p_i) em relação a chance contra a ocorrência de um evento de ravina ($1-p_i$), conforme a Equação 1:

$$OR = \frac{p_i}{1 - p_i} \quad (1)$$

Os valores OR foram plotados em relação aos valores do ponto médio de cada uma das 10 classes dos mapas VI, LD, SD e RD. Em seguida, os valores originais destes mapas foram transformados em valores fuzzy, utilizando-se as seguintes funções de afinidade (Mf): mapa VI - função sigmoideal decrescente; mapa RD - função linear crescente; mapa LD - função triangular; mapa SL - função Gaussiana. O risco de erosão no pixel i (R_i) foi calculado por meio de um modelo baseado na soma ponderada dos valores fuzzificados dos mapas VI, SL, LD e RD no pixel i , de acordo com a Equação 2:

$$R_i = F_{OR-SL} \cdot SL_i + F_{OR-RD} \cdot RD_i + F_{OR-LD} \cdot LD_i + F_{OR-VI} \cdot VI_i \quad (2)$$

onde, F_{OR-LD} , F_{OR-RD} , F_{OR-VI} são, respectivamente, os coeficientes de regressão logística entre OR e os valores originais dos mapas SL, LD, RD e VI; SL_i , LD_i , RD_i and VI_i são, respectivamente, os valores fuzzificados de SL, SD, RD e IV no pixel i .

Resultados e Discussão

A Figura 2 mostra, como exemplo, o diagrama relacionando os valores de OR (a chance em favor de um evento de ravina em relação à chance contrária a ele), em função do valor do centro médio das classes do mapa de índice de vegetação. A Figura 3 mostra o mapa de risco à ravinação na APA Fernão Dias, gerado por meio de operação espacial baseada em álgebra de mapas, de acordo com a Equação 2.

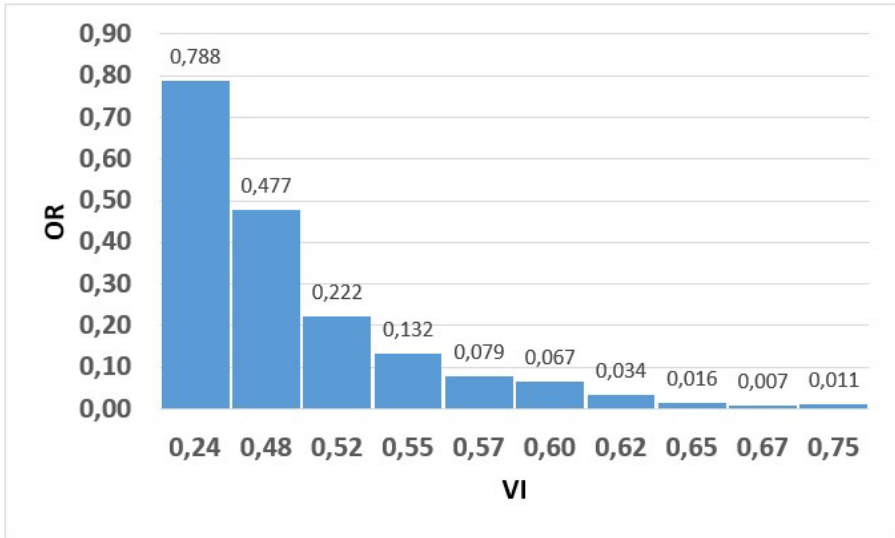


Figura 2 - Diagrama representando a distribuição dos valores de OR em relação ao centro médio das classes do mapa de IV.

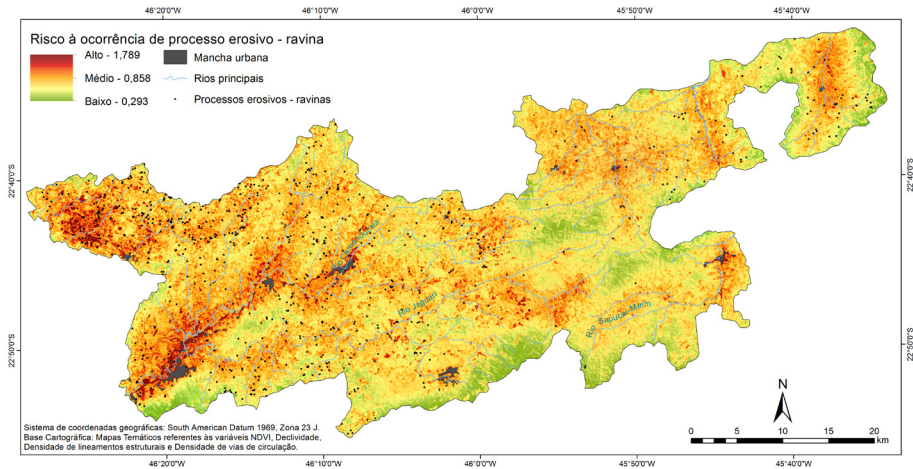


Figura 3 - Mapa de risco à ocorrência de ravinas na APA Fernão Dias - MG, calculado a partir do modelo representado pela Equação 2.

Conclusões

A metodologia proposta neste artigo apresenta soluções para minimizar problemas associados à atribuição de pesos espaciais a variáveis utilizadas na identificação de áreas sob risco de erosão, e também, ao mapeamento destas áreas. A primeira solução refere-se a estimativa do peso de cada mapa temático envolvido no risco à erosão. Para estimar estes pesos utilizamos valores do coeficiente de regressão logística calculados entre a variável dicotômica (presença ou ausência

de ravinas) e os valores dos mapas das variáveis ambientais que contribuem para a ocorrência de ravinas. Como estes coeficientes foram calculados a partir de dados reais obtidos de mapeamento de ravinas existentes no terreno, os valores dos pesos atribuídos aos mapas foram mais precisos que aqueles utilizados em procedimentos convencionais subjetivos.

A segunda solução apresentada neste artigo baseia-se no uso da lógica de decisão fuzzy para atribuir a cada pixel a um determinado grau de risco à erosão. Neste caso, funções de afinidade foram atribuídas individualmente a cada variável ambiental, de acordo com curvas de dependência logística entre a chance de ocorrência de ravinas e os valores dos pixels dos mapas de índice de vegetação, densidade de estradas, densidade de lineamentos estruturais e declividade do terreno. Este procedimento permitiu um ajuste mais realista entre os valores dos pixels dos mapas e sua associação a um grau de risco de erosão.

Baseado na variabilidade dos valores de OR estimados entre a ocorrência/não ocorrência de ravinas e os valores dos mapas ambientais, no território da APA Fernão Dias, observamos que estas variáveis ambientais responderam diferentemente em relação a ocorrência de ravinas. A chance de ocorrer ravina foi muito elevada em áreas com baixo índice de vegetação (OR=0.788 em VI=0.24), e muito baixa em áreas com alto índice de vegetação (OR=0.001 at VI=0.75). No caso da densidade de estradas a dependência foi inversa, isto é, baixa chance de ocorrência de ravinas em áreas com baixa densidade de estradas (OR=0.043 em RD=0.40 km/km²) e alta chance em áreas com alta densidade de estradas (OR=0.169 em RD=2.71 km/km²). As demais variáveis - densidade de lineamentos estruturais e declividade do terreno mostraram, em relação às anteriores, relativamente menor influência na chance de ocorrência de ravinas. Em síntese, os graus de associação entre as variáveis ambientais estudadas e a chance de ocorrer ravinas podem assim ser ordenados: VI (R² = 0,8958); RD (R² = 0,6573); LD (R² = 0,225); SL (R² = 0,072).

A menor influência da declividade do terreno na ocorrência de ravinas se deu, principalmente, pelo fato de as áreas mais declivosas estarem protegidas legalmente dentro da APA, com restrição ao uso agropecuário e cobertas com remanescentes florestais, correspondendo a altos valores de IV. As áreas com topografia mais plana, por outro lado, encontram-se ocupadas com maior densidade de atividades agropecuárias e urbanas.

O modelo espacial aqui apresentado deve ser utilizado sob determinadas condições específicas. Os valores dos pesos e a equação de cálculo do risco são intrínsecos às características físico-geográficas de cada área a ser mapeada, e podem variar a depender das configurações geomorfológica, geológica e de uso e cobertura do solo da região. A maior eficiência do uso deste modelo está em condições em que o pesquisador deseja identificar locais que estarão sob risco futuro de ravinamento, com base em evidências já observadas em outros locais da mesma área de estudo. Portanto, o modelo utiliza a associação entre a ocorrência do fenômeno e a estrutura espacial observada e confirmada a priori, para simular e mapear, a posteriori, estas mesmas relações em outros locais da mesma área. Este modelo é adequado também ao mapeamento de riscos em áreas onde não haja a disponibilidade de mapas pedológicos ou geológicos em escalas médias a grandes.

Apoio Financeiro - FAPESP e CAPES.

Bibliografia

- Burrough, P.A. Fuzzy mathematical methods for soil survey and land evaluation. *Journal of Soil Science* 40: 477-492, 1989.
- Burrough, P. A.; MacMillan, R. A., and van Deursen, W. 1992. Fuzzy classification methods for determining land suitability from soil profile observations and topography. *Journal of Soil Science* 43: 193-210.
- ESRI - ArcGIS 10.0. ESRI, Redlands, California, USA, 2010.
- Garófalo, D. F. T.; Ferreira, M. C.; Ferreira, M. F. M.; Messias, C. G. - Mapeamento da densidade de processos erosivos e de movimentos de massa, a partir de imagens Google Earth georreferenciadas, utilizando o estimador kernel: uma aplicação na APA Fernão Dias (MG). *Anais do IX Simpósio Nacional de Geomorfologia*, Rio de Janeiro, 2012.
- Panik, M. - *Regression modeling: methods, theory and computation with SAS*. CRC Press, Boca Raton, FL, USA, 2009.
- Zadeh. L.A. (1965) - Fuzzy sets. *Information control*, 8:338-353.
- Wang, F.; Hall, G.B. - Fuzzy representation of geographical boundaries in GIS. *International Journal of Geographical Information Systems*, 10(5):573-590, 1996.