

AVALIAÇÃO DA SUSCEPTIBILIDADE A DESLIZAMENTOS ROTACIONAIS ATRAVÉS DA APLICAÇÃO DE MÉTODOS ESTATÍSTICOS

Sérgio Cruz de Oliveira¹; Jorge Rocha²; José Luís Zêzere¹; Ricardo A. C. Garcia¹; Aldina Piedade¹

¹RISKam – Centro de Estudos Geográficos da Universidade de Lisboa – Faculdade de Letras, Alameda da Universidade, 1600-214 Lisboa

²MOPT – Centro de Estudos Geográficos da Universidade de Lisboa – Faculdade de Letras, Alameda da Universidade, 1600-214 Lisboa

(cruzdeoliveira@fl.ul.pt; jrocha@fl.ul.pt; jlzezere@fl.ul.pt; rgarcia@fl.ul.pt; ald.piedade@fl.ul.pt)

RESUMO

Grande parte do esforço científico na definição dos mapas de susceptibilidade à ocorrência de deslizamentos prende-se sobretudo com o método estatístico utilizado na avaliação de susceptibilidade e respectiva validação dos modelos preditivos. Neste trabalho utilizou-se um conjunto de métodos estatísticos (Valor Informativo, *Weights of Evidence*, Probabilidade *Bayesiana*, Lógica Difusa, Regressão Logística, Distância de *Mahalanobis*, Máxima Entropia e Algoritmo Genético – GARP), numa área a norte de Lisboa, de forma a perceber qual a influência que a mudança de método estatístico exerce nos mapas de susceptibilidade. Como variáveis independentes usaram-se seis temas (declive, exposição, litologia, uso do solo, tipo de solo e curvatura das vertentes) e como variável dependente um inventário de 64 deslizamentos rotacionais. A unidade de terreno definida no processo de computação foi o *pixel* (5 m) e a entrada dos deslizamentos nos modelos fez-se quer através de um único ponto por movimento (centróide), quer com a totalidade da área deslizada. Embora se tenham verificado algumas diferenças nos mapas de susceptibilidade fruto dos diferentes métodos utilizados, os resultados obtidos permitem concluir que os modelos preditivos são mais sensíveis à forma como a variável dependente entra nos modelos (pontos vs área) do que ao tipo de método estatístico.

1. Introdução

A avaliação da susceptibilidade de ocorrência de deslizamentos, à escala regional recorre, tradicionalmente, a dois tipos de métodos: métodos directos (baseados em critérios geomorfológicos) e métodos indirectos (de base estatística e probabilística). Actualmente, grande parte do esforço científico na definição dos mapas de susceptibilidade a deslizamentos prende-se com o tipo de unidade de terreno utilizada na modelação, o método estatístico de avaliação de susceptibilidade e a validação dos modelos preditivos. Neste sentido, importa salientar que as relações entre deslizamentos, vertentes, e todos os factores que condicionam a sua evolução são frequentemente complexas, multi-factoriais e de difícil descrição matemática (Lu e Rosenbaum, 2003). Acresce a este facto a dificuldade

em caracterizar espacialmente as variáveis de base física na escala regional, aumentando por isso o desafio de predição de novas áreas instáveis (Lu e Rosenbaum, 2003). Estas condicionantes conduziram a uma expansão de abordagens estatísticas/probabilísticas indirectas, de base quantitativa. Contrariamente aos métodos heurísticos, onde os factores são ponderados de acordo com indexações geralmente semi-quantitativas, nas abordagens probabilísticas a importância dos factores ditos condicionantes (ou de predisposição) da instabilidade é função da sua relação estatística com a distribuição dos deslizamentos já ocorridos no território em estudo. Nestes casos, a avaliação da susceptibilidade é baseada na assunção de que as condições que geraram instabilidade de um determinado tipo no passado serão as mesmas que originarão novas áreas instáveis do mesmo tipo no futuro (Carrara *et al.*, 1999). Desta forma, torna-se possível prever a distribuição futura da instabilidade nas vertentes pela correlação entre a distribuição espacial dos deslizamentos que ocorreram no passado e um conjunto de variáveis independentes que representam os factores condicionantes da instabilidade (por exemplo, declive, litologia, uso do solo, etc). No contexto da aplicação de métodos de base estatística, a grande dúvida, manifestada pela quantidade de literatura produzida, reporta-se à metodologia que mais fidedignamente caracterize a susceptibilidade a deslizamentos de um determinado tipo. Deste modo, este estudo tem como principais objectivos: (i) avaliar a susceptibilidade aos deslizamentos rotacionais, numa única área de estudo, através de vários métodos estatísticos/probabilísticos (por exemplo, Valor Informativo, *Weights of Evidence*, Probabilidade Bayesiana, Regressão Logística, Lógica Difusa, Distância de *Mahalanobis*, Máxima Entropia e Algoritmo Genético GARP (*Genetic Algorithm for Rule-set Production*)); (ii) definir a sensibilidade dos mapas de susceptibilidade à aplicação dos diferentes métodos estatísticos; (iii) comparar os resultados com base em técnicas de validação cruzada, computação das taxas de sucesso e de predição e cálculo das respectivas áreas abaixo da curva (AUC – *Area Under the Curve*).

2. Enquadramento da área de estudo

A área de estudo localizada na região a norte de Lisboa, com cerca de 12,5 km², abrange duas pequenas bacias hidrográficas (Rio Salema e Ribeira da Laje), afluentes do Rio Grande da Pipa, Arruda dos Vinhos (Figura 1). A altitude varia entre os 65 m e os 370 m, e o substrato é dominado pela presença de calcários compactos, margas, argilas e arenitos datados do Kimeridgiano, Jurássico Superior. Estruturalmente, a área de estudo encontra-se no flanco sul de um empolamento tectónico regional, que deu origem a uma deformação anticlinal de grande raio de curvatura centrada em Arruda dos Vinhos. Deste modo, as formações geológicas na área de estudo encontram-se inclinadas de uma forma geral para sul, com pendor próximo dos 5°. Do ponto de vista geomorfológico, predominam as vertentes anaclinais (contrárias à inclinação das camadas), de perfil côncavo, devido à presença das margas, argilas e arenitos que compõem a formação das Margas da

Abadia, dominadas por cornijas talhadas nos calcários corácicos de Amaral. Neste contexto geomorfológico particular, a instabilidade de vertentes é marcada pelo predomínio de deslizamentos associados a superfícies de ruptura do tipo rotacional (profunda e superficial) e deslizamentos associados a rupturas translacionais superficiais.

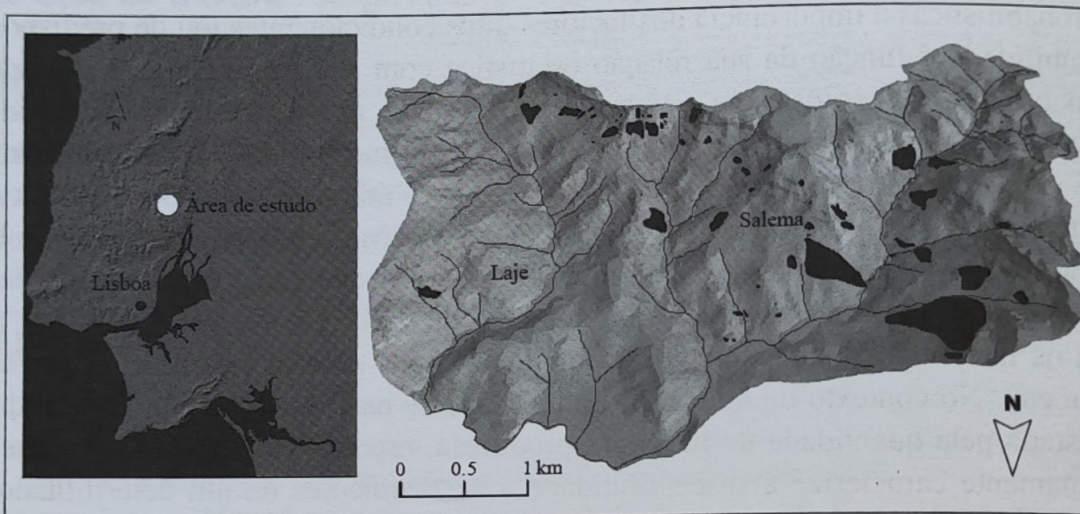


Figura 1 – Localização da área de estudo (bacias do Rio Salema e Ribeira da Laje, Arruda dos Vinhos) e dos deslizamentos rotacionais (polígonos a negro).

O inventário dos deslizamentos foi obtido com recurso a trabalho de campo, com cartografia apoiada sobre ortofotomapas na escala 1:2000. Foram inventariados cerca de 150 movimentos de vertente dos quais se destacam, 64 deslizamentos rotacionais que representam 97,3% da área total instabilizada e cerca de 4% da área de estudo. Os deslizamentos rotacionais têm uma densidade aproximada de 5 mov/km² e a sua importância ficou demonstrada no passado em função do potencial destruidor que lhes está associado, razão pela qual foram seleccionados neste trabalho.

3. Metodologia

3.1 Avaliação da susceptibilidade: aquisição e gestão de dados

Na avaliação da susceptibilidade foram usadas 6 variáveis independentes (uso do solo, declive, exposição, perfil transversal das vertentes, litologia, tipo de solo), com 49 classes previamente definidas. Destas variáveis 3 derivaram, directa ou indirectamente, do modelo digital de terreno (declive, exposição e perfil transversal das vertentes) construído a partir da informação disponível na escala 1:25 000, com uma equidistância entre curvas de nível de 10 m. Como variável dependente, foi utilizado o inventário dos deslizamentos rotacionais (64 casos). A unidade de terreno definida no processo de computação foi uma célula de 5x5 m² e a entrada dos deslizamentos nos modelos fez-se quer através de um

único ponto (centróide) por movimento (métodos dos *Weights of Evidence* e Regressão Logística, por constrangimento de *software*), quer considerando a totalidade da área deslizada (restantes métodos). A Figura 2 esquematiza a abordagem metodológica utilizada na avaliação da susceptibilidade a deslizamentos rotacionais.

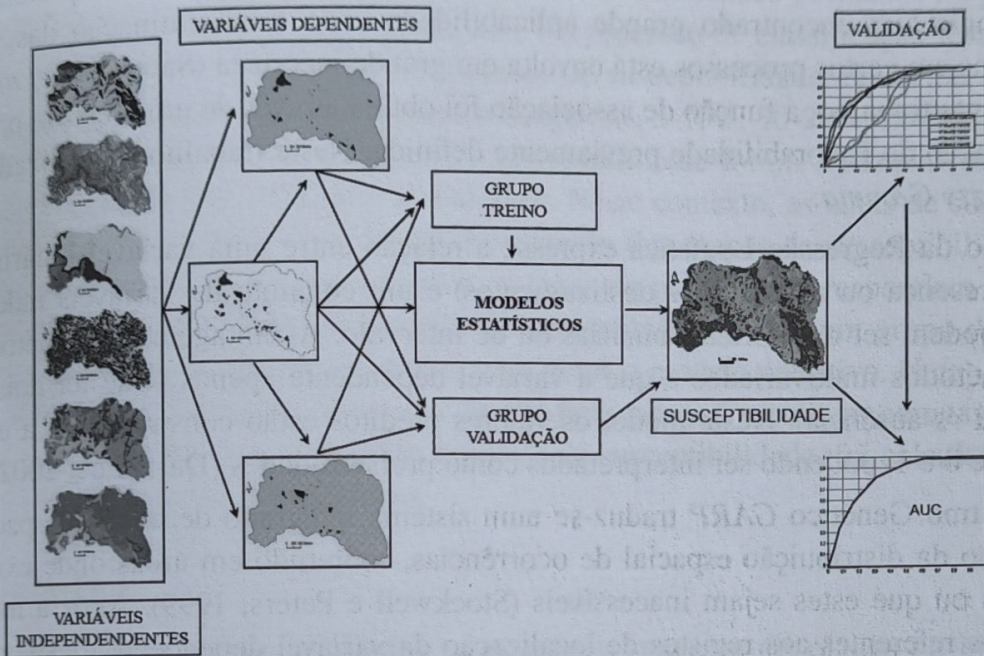


Figura 2 – Metodologia de avaliação da susceptibilidade a deslizamentos rotacionais.

3.2 Métodos estatísticos

O Valor Informativo (VI) é um método que descreve quantitativamente a relação entre cada variável independente e a variável dependente (deslizamentos) sob a forma de uma pontuação. O valor final de susceptibilidade para um determinado *pixel* é dado pela soma do VI obtido para cada variável. Quanto maior o valor final do VI, maior a predisposição para a instabilidade nessa unidade de terreno (Yin e Yan, 1988).

No teorema de Bayes, o logaritmo da FDP (Função Densidade de Probabilidade) posterior é igual à soma do logaritmo da FDP da amostra com o logaritmo da FDP anterior (*prior*) com o logaritmo inverso da FDP preditiva da amostra. Neste sentido, a Probabilidade *Bayesiana* (PB) incorpora o conceito de Função de Favorabilidade (Chung e Fabbri, 1993), em que a predição da susceptibilidade pode ser considerada como a probabilidade condicional conjunta de uma determinada unidade de terreno ser afectada por um deslizamento futuro, condicionado pelas características físicas da área (Zêzere *et al.*, 2008).

No caso dos *Weights of Evidence* (WofE), Bonham-Carter *et al.* (1989) demonstraram que o logaritmo da probabilidade posterior de uma determinada ocorrência numa deter-

minada unidade de terreno pode ser definido pela adição de um coeficiente (*Weight*) por cada mapa introduzido nos logaritmos das probabilidades *a priori*. O produto final é um mapa das probabilidades posteriores.

Na Lógica Difusa (LD) é especificada uma função de pertença que varia entre 0 e 1 permitindo quantificar a importância de uma determinada variável (grau de pertença). Esta metodologia tem encontrado grande aplicabilidade onde a determinação das variáveis envolvidas em certos processos está envolta em grande incerteza (Saboya Jr. *et al.*, 2006). No presente trabalho, a função de associação foi obtida através de uma relação proporcional à Função de Favorabilidade previamente definida. Neste trabalho foi utilizado o operador *Fuzzy Gamma*.

O método da Regressão Logística expressa a relação entre uma variável binária dependente (presença ou ausência de deslizamentos) e um conjunto de variáveis independentes que podem ser categóricas, binárias ou de intervalo. A vantagem deste método sobre outros métodos multivariados é que a variável dependente apenas pode ter dois valores (presença *vs* ausência). Deste modo, os valores preditos estão restringidos a um intervalo entre 0 e 1, podendo ser interpretados como probabilidades (Dai e Lee, 2002).

O Algoritmo Genético *GARP* traduz-se num sistema integrado de análise espacial para a predição da distribuição espacial de ocorrências, sobretudo em áreas onde exista falta de dados ou que estes sejam inacessíveis (Stockwell e Peters, 1999). Após a integração dos dados referentes aos registos de localização da variável dependente (deslizamentos), estes são analisados com recurso a procedimentos analíticos baseados em auto-aprendizagem. O sistema pode tornar-se mais robusto através da utilização de algoritmos genéticos, árvores de decisão e redes neuronais.

No método da Máxima Entropia (Maxent) estima-se a distribuição alvo através da geração da distribuição de máxima entropia (*i.e.*, mais próximo do uniforme), sujeita ao constrangimento de que o valor esperado para cada evento sob esta distribuição estimada é equivalente à sua média empírica. As vantagens da utilização do Maxent passam pela possibilidade de utilizar dados contínuos ou categóricos, bem como interações entre as diferentes variáveis, possibilitando a construção de algoritmos determinísticos de forma a garantir a convergência à óptima distribuição probabilística, logo máxima entropia (Phillips *et al.*, 2006).

A distância de *Mahalanobis* corresponde a uma medida de afastamento *n* dimensional a um valor médio de um conjunto de dados. Esta medida, quando aplicada à modelação da susceptibilidade de ocorrência de deslizamentos, calcula a similaridade existente entre o valor multidimensional médio do conjunto de amostras de ocorrência e cada um dos locais da área de estudo considerada. O cálculo da distância de *Mahalanobis* envolve sempre duas variáveis, um vector de médias e uma matriz de covariâncias e tem em consideração a dispersão dos dados e a orientação das amostras.

3.3 Classificação e validação

Neste trabalho foram utilizadas 3 bases de dados de deslizamentos. Uma primeira com os 64 deslizamentos rotacionais e mais duas obtidas a partir da partição aleatória da base inicial, cada uma com 32 deslizamentos (uma utilizada como grupo de treino e outra como grupo de validação). Para cada um dos métodos estatísticos utilizados foi produzido um mapa de susceptibilidade não classificado. No processo de classificação foram definidas as seguintes classes, por ordem decrescente de susceptibilidade, de forma a permitir a comparação dos resultados: os 5% mais susceptíveis, 5-10%, 10-20%, 20-30%, 30-40% e 40%-100%. Na validação dos mapas de susceptibilidade foram utilizadas técnicas de validação cruzada (*cross-validation techniques*). Neste contexto, as taxas de sucesso correspondem às curvas obtidas através do cruzamento dos mapas de susceptibilidade com os deslizamentos que lhe deram origem e as taxas de predição correspondem às curvas obtidas através do cruzamento dos mapas de susceptibilidade com um grupo de deslizamentos independente do grupo que serviu para gerar o modelo preditivo. Uma vez que as curvas referentes às Taxas de Sucesso e de Predição apenas permitem comparar visualmente as diferenças entre os diferentes modelos de susceptibilidade, foi calculada a AUC de forma a quantificar essas diferenças.

4. Resultados

Numa primeira fase foram escolhidos os principais métodos abordados na bibliografia (PB, VI, FL, RL e WofE) e foram produzidos os respectivos mapas de susceptibilidade (Figura 3), computadas as correspondentes curvas de sucesso (Figura 4) e calculadas as AUC (Tabela 1).

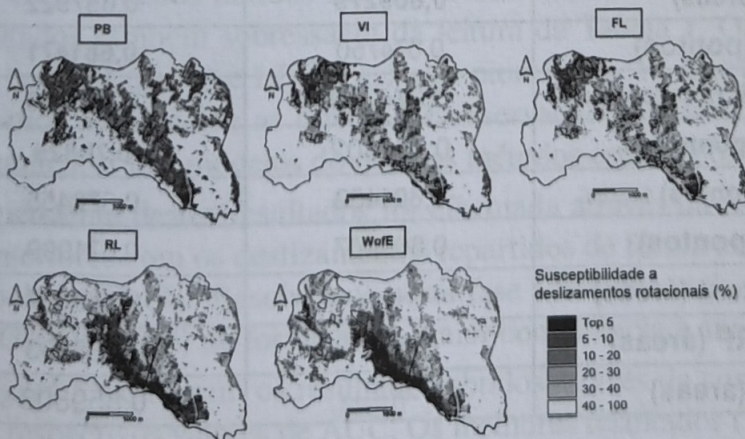


Figura 3 – Mapas de susceptibilidade a deslizamentos rotacionais obtidos pelos seguintes métodos estatísticos: PB, VI, LD (operador Gamma), RL e WofE.

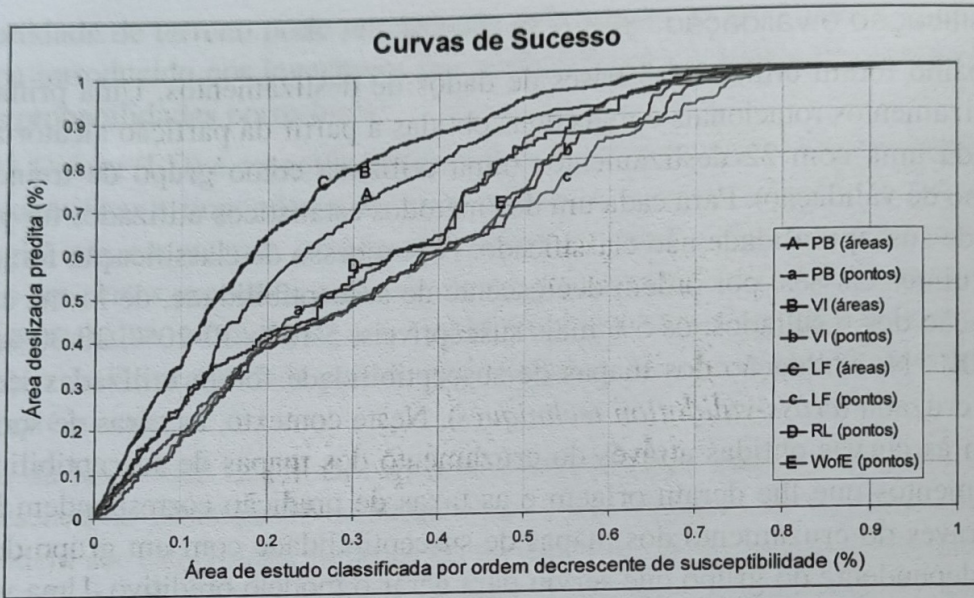


Figura 4 – Curvas de Sucesso obtidas para os seguintes métodos estatísticos: PB, VI, LF, RL e WofE. Nas curvas A, B, C os deslizamentos entraram nos modelos com a totalidade da área deslizada. Nas curvas a, b, c, D, E os deslizamentos entraram como um único ponto (um pixel por deslizamento). A validação foi em ambos os casos feita com a totalidade da área deslizada.

Tabela 1 – Valor das AUC calculadas para as Curvas de Sucesso e de Predição.

	Taxas de Sucesso (todos)	Taxas de Predição (Grupo A)
PB (áreas)	0,758854	0,661232
VI (áreas)	0,808437	0,656235
LF (áreas)	0,809279	0,657922
RL (pontos)	0,674750	0,681471
WofE (pontos)	0,674750	0,693380
PB (pontos)	0,701802	0,70829
VI (pontos)	0,601452	0,672455
LF (pontos)	0,660407	0,671099
Maxent (áreas)	-----	0,668057
GARP (áreas)	-----	0,588833
DM (áreas)	-----	0,559803

Da interpretação da Figura 3 sobressaem dois padrões de mapas de susceptibilidade. Um definido pelos métodos PB, VI e FL em que as áreas classificadas como mais susceptíveis se encontram mais dispersas pela área de estudo e um outro definido pelos métodos RL e

WofE em que as áreas mais susceptíveis se encontram concentradas espacialmente. Relativamente a este último padrão, ele prediz mal os movimentos de maiores dimensões (por exemplo, deslizamento da Capelã, o maior deslizamento da área de estudo). Este facto encontra-se bem evidenciado na Figura 3, se atentarmos aos 10% da área classificada como mais susceptível.

A Figura 4 reporta-se à capacidade preditiva dos mapas de susceptibilidade expressa através das Curvas de Sucesso. Se considerarmos os cinco métodos que deram origem aos mapas da Figura 3 (PB, VI, FL, RL e WofE), curvas A, B, C, D, E, respectivamente, podemos observar que os métodos RL e WofE são os que apresentam piores resultados. Se aos 10% de área classificada como mais susceptível as diferenças não são tão evidentes, já aos 30% estas aumentam consideravelmente (20% a 30% de diferenças da capacidade preditiva).

O desempenho menos interessante dos métodos RL e WofE deve ser interpretado com prudência, uma vez que a forma de entrada da variável dependente nestes modelos é diferente dos restantes. Com efeito, nos modelos de susceptibilidade feitos com base na RL e WofE, os deslizamentos entraram apenas com um único ponto por movimento, ou seja, 64 pontos (64 *pixels*), enquanto que nos restantes modelos entraram com a totalidade da área deslizada (19 970 *pixels*). Desta forma, é de admitir que a utilização de apenas um ponto por movimento seja suficiente para captar as características físicas dos movimentos de menores dimensões, tendo mais dificuldade em caracterizar os movimentos maiores.

A ser verdade que os mapas de susceptibilidade perdem capacidade preditiva se os deslizamentos entrarem apenas com um único ponto por movimento, esse facto deve ser demonstrado aplicando os modelos PB, VI e FL com pontos e não com áreas (curvas a, b, c na Figura 4). Os resultados obtidos são muito semelhantes aos gerados com a RL e WofE. Estes resultados também sobressaem da leitura da Tabela 1. Os melhores modelos foram obtidos com VI (áreas) e LF (áreas) e os piores com VI (pontos) e LF (pontos). Desta forma podemos afirmar que as diferenças observadas se devem mais ao modo de entrada dos deslizamentos nos modelos do que aos métodos estatísticos utilizados.

A validação independente destes resultados foi efectuada através da repetição das corridas de modelos anteriores com os deslizamentos repartidos de forma aleatória num grupo de treino e noutro de validação. Resolvemos nesta fase adicionar mais três métodos estatísticos (Maxent, GARP e DM) de forma a dar maior consistência à análise.

A Tabela 1 e a Figura 5 expressam os resultados obtidos através da computação das Curvas de Predição e respectivos valores de AUC. Os melhores resultados (Tabela 1) são obtidos com a PB (pontos), WofE (pontos) e RL (pontos) e os piores com DM (áreas) e GARP (áreas). Também aqui se evidenciam os dois padrões descritos anteriormente (Figura 5). As curvas respeitantes aos métodos RL e WofE, apresentam melhores resultados de pre-

dição praticamente até aos 40% de área classificada como mais susceptível. Os restantes métodos apenas se aproximam da tendência destas curvas depois dos 40% de área classificada como mais susceptível.

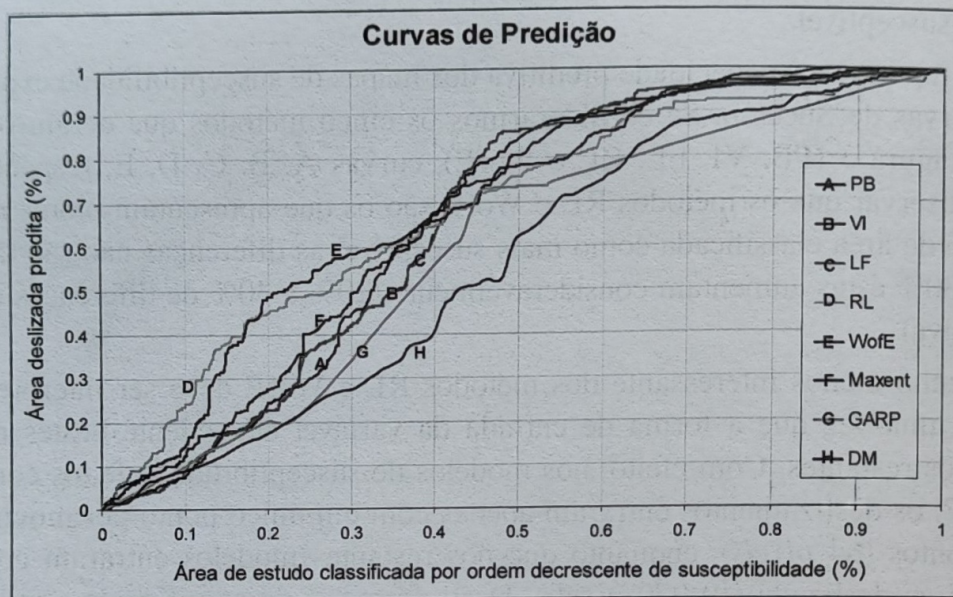


Figura 5 – Curvas de Predição obtidas para todos os métodos estatísticos considerando dois grupos de deslizamentos independentes, um para treino (gera os mapas de susceptibilidade) e um outro para validação.

No entanto, importa referir que os resultados relativos à capacidade preditiva dos mapas de susceptibilidade a deslizamentos rotacionais feita de forma independente não se mostraram satisfatórios, uma vez que temos que isolar 40% da área classificada como mais susceptível para prever entre 60% a 70% dos novos movimentos.

5. Conclusões

A utilização dos diferentes métodos estatísticos não permite tirar uma conclusão definitiva sobre qual ou quais os melhores métodos para avaliar a susceptibilidade a deslizamentos rotacionais. Os melhores resultados em termos de sucesso obtiveram-se com VI e LF, enquanto que na predição os métodos PB, WofE e RL revelaram-se mais eficazes. No entanto, as diferenças observadas num caso e noutro não são muito significativas e devem ser lidas com cautela, atendendo à dimensão reduzida da amostra correspondente à variável dependente. Este trabalho permitiu evidenciar, contudo, dois padrões de mapas de susceptibilidade, mais sensíveis à forma de entrada da variável dependente (deslizamentos como pontos ou com a totalidade da área deslizada) do que aos métodos estatísticos.

Agradecimentos

Este estudo é parte do Projecto *MapRisk* (PTDC/GEO/68227/2006) financiado pela Fundação para a Ciência e a Tecnologia. Agradecemos ao Prof. Doutor João Casaca os comentários e sugestões ao trabalho.

Referências Bibliográficas

- Bonham-Carter, G. F.; Agterberg, F. P.; Wright, D. F. (1989). "Weights of evidence modeling: a new approach to mapping mineral potential". In: Agterberg, F. P. and Bonham-Carter, G. F. (eds.) *Statistical applications in the Earth Science, Geological Survey of Canada Paper 89-9*, Ottawa, Canada, Págs. 171-183.
- Carrara, A.; Guzzetti, F.; Cardinali, M.; Reichenbach, P. (1999). "Use of GIS Technology in the Prediction and Monitoring of Landslide Hazard". *Natural Hazards*, Vol. 20, No. 2-3, Págs. 117-135.
- Chung, C. F.; Fabbri, A.G. (1993). "The representation of geoscience information for data integration". *Non-renewable Resources*, Vol. 2, No. 2, Págs. 122-139.
- Dai, F. C.; Lee, C. F. (2002). "Landslide characteristics and slope instability modelling using GIS, Lantau Island, Hong Kong". *Geomorphology*, No. 42, Págs. 213-228.
- Lu, P.; Rosembaum, M. S. (2003). "Artificial Neural Networks and Grey Systems for the prediction of slope stability". *Natural Hazards*, No. 30, Págs. 383-398.
- Phillips, S. J.; Anderson, R. P.; Schapire, R. E. (2006). "Maximum entropy modeling of species geographic distribution". *Ecological Modelling*, No. 190, Págs. 231-259.
- Saboya Jr., F.; Alves, M. G.; Pinto, W. D. (2006). "Assessment of Failure Susceptibility of Soil Slopes Using Fuzzy Logic". *Engineering Geology*, Vol. 86, Págs. 211-224.
- Stockwell, D.; Peters, D. (1999). "The GARP modelling system: problems and solutions to automated spatial prediction". *International Journal of Geographical Information Science*, Vol. 13, No. 2, Págs. 143-158.
- Yin, K. L.; Yan, T. Z. (1988). "Statistical prediction model for slope instability of metamorphosed rocks". In: Bonnard, C. (Ed.), *Proceedings 5th International Symposium on Landslides*, Lausanne, Switzerland, Vol. 2, Págs. 1269-1272.
- Zêzere, J. L.; Garcia, R. A. C.; Oliveira, S. C.; Reis, E. (2008). "Probabilistic landslide risk analysis considering direct costs in the area north of Lisbon (Portugal)". *Geomorphology*, Vol. 94, No. 3-4, Págs. 467-495.