

A importância do processo de classificação de dados na cartografia: um exemplo na cartografia de susceptibilidade a movimentos de vertente.

The classification data process in cartography: an example on landslide susceptibility maps.

R. A. C. Garcia^{1,2}, J. L. Zêzere¹, S. Cruz de Oliveira^{1,2}

¹ Centro de Estudos Geográficos, Universidade de Lisboa, racgarcia@yahoo.com, jlzezere@fl.ul.pt, cruzdeoliveira@gmail.com;

² Bolseiro da Fundação para a Ciência e a Tecnologia, Ministério da Ciência, Tecnologia e Ensino Superior.

Resumo: A expressão gráfica dos resultados de uma investigação é sem dúvida um dos passos mais importantes num estudo, pois é através de mapas que muitos dos resultados são divulgados e utilizados, nomeadamente no Ordenamento do Território. Apesar de muitas vezes ignorado, sobretudo com a automatização de processos que os Sistemas de Informação Geográfica permitem, o método de classificação de dados e número de classes utilizado para a elaboração de um mapa pode trazer à visualização dos resultados diferenças consideráveis. Com o intuito de verificar as desigualdades entre métodos de divisão em classes e de desenvolver um índice que permita avaliar qual o mais adequado a um objectivo específico, realizou-se um estudo comparativo da susceptibilidade a deslizamentos rotacionais na área da Bacia da Abadia (Torres Vedras, Portugal). Através dos resultados obtidos comprova-se que o processo de divisão em classes tem uma importância muito significativa na expressão cartográfica das avaliações de susceptibilidade, podendo-se elaborar mapas completamente diferentes e formalmente correctos, bastando para tal modificar o método de classificação de dados. O processo baseado nas taxas de sucesso é o método que apresenta resultados melhor adaptados à realidade, não obstante ser uma metodologia que se reveste de alguma subjectividade. Qualquer resultado obtido deve ser validado e criticamente avaliado, pelo que o conhecimento da área de estudo se torna imprescindível. Deste modo, fica provado que esta etapa de trabalho deve ser tida em conta e não realizada de forma displicente, uma vez que da sua correcta execução dependerá a boa interpretação dos resultados, podendo uma má opção na escolha da metodologia de divisão em classes fazer com que todo o trabalho realizado tenha sido em vão.

Palavras-chave: Classificação de dados, deslizamentos, mapas de susceptibilidade, taxas de sucesso.

Abstract: The graphical expression of the results obtained in research is one of the most important steps because many of them are represented by maps, mainly when will be used on territorial management. Nevertheless, be frequently mislaid mistreated, mostly because of the automatic procedures which Geographical Information Systems allows the classification data process and the number of classes used can take to the maps considerable differences. With the aim of identify the differences between types of classification data methods and develop an index that permit to evaluate which of them is the better one to serve our goals, a susceptibility to rotational landslides, in the Abadia Basin (Torres Vedras, Portugal), comparative study was done. By the results obtained it was proved that the classification of data in classes have an important influence on the susceptibility assessment research outcome, because with the same data base it's possible create several complete different maps only using and changing standard and acceptable data classification processes. The classification based on the success rates it was the one that had the better results of the capacity index and the map is the most adjusted to reality. However, each result should be validated and a critical analysis has to be done based on a indispensable field work and knowledge. With this work it was proved the importance of the classification data stage which shouldn't be done negligently because from the correct execution of that step depends the good interpretation of the results. A bad choice of the methodology used could obliterate all the previous research work.

Keywords: Data classification, landslides, susceptibility maps, success rates.

1. INTRODUÇÃO

A representação gráfica dos resultados de uma investigação é uma etapa de elevada importância na elaboração de trabalhos com base cartográfica. A eficiência da divulgação dos resultados e a sua

utilização, nomeadamente no Ordenamento do Território ou com fins pedagógicos, depende muito das opções tomadas durante esta fase. Esta questão é abordada neste trabalho no contexto da produção de mapas de susceptibilidade a movimentos de vertente.

Como se pode observar na figura 1, onde se representam, de modo simplificado, as fases de elaboração de um mapa de susceptibilidade, são diversas as etapas que condicionam os resultados finais.

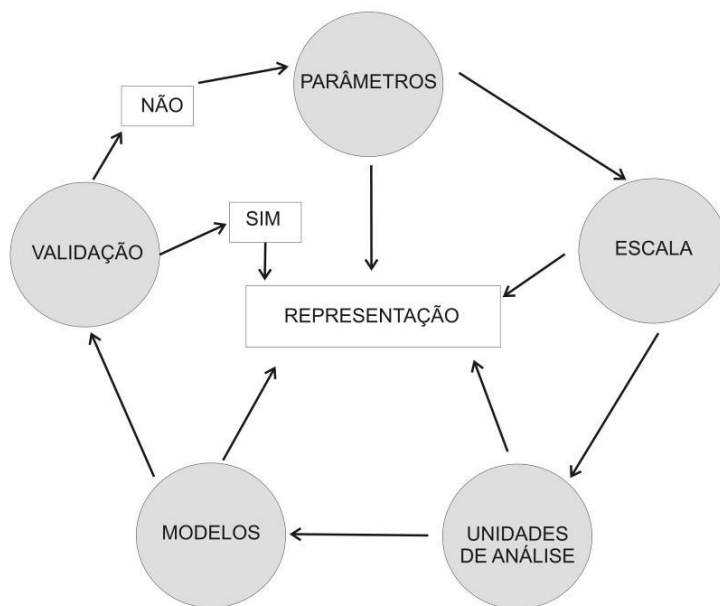


Figura 1 – Etapas de elaboração de um mapa de susceptibilidade.

Figure 1 – Steps for execution of susceptibility maps.

A escolha dos parâmetros a utilizar é em geral a primeira fase de um estudo, a que se segue o grau de pormenor a que será executado, com a escolha das escalas a usar. A definição das unidades de análise é a terceira etapa, sendo a sua definição baseada em critérios geomorfológicos (p.e.: cruzamento entre declives e litologia, sub-bacias hidrográficas, etc.) ou automáticos (p.e.: pixel). A definição das unidades de trabalho pode ser condicionada pela tipologia de técnicas que se vão utilizar, podendo estas ser de avaliação absoluta (geotécnicas, determinísticas), baseada nos parâmetros e relações físicas do terreno, ou de avaliação relativa (geomorfológicas), assentes na distribuição de movimentos e seu relacionamento com os factores condicionantes. A escolha destes métodos depende não só do objectivo final, mas das técnicas, meios e fundos que o investigador domina e tem ao seu dispor (Carrara, 1993). Os modelos relativos podem apoiar-se em métodos de avaliação directa (cartografia directa), fortemente condicionada pela experiência do investigador e por tal considerados de elevada subjectividade, ou em métodos de avaliação indirecta (indexação, análise estatística bivariada ou multivariada), relacionando de forma mais ou menos automática instabilidades e factores condicionantes, com a assunção de que os factores que geraram a instabilidade passada irão produzir os mesmos efeitos no futuro. Estes métodos têm a grande vantagem de poderem ser avaliados e validados logo após a sua elaboração, através de partições temporais, espaciais ou aleatórias das bases de dados, e não apenas aguardando pela ocorrência de novos eventos, como acontece com os métodos de cartografia directa. Consoante os resultados da

validação prossegue-se para o processo de representação ou ir-se-ão reformular uma ou várias das etapas anteriores, até que os resultados estejam de acordo com os objectivos a atingir.

A representação dos resultados é a última fase do trabalho, estando esta interligada e dependente de todas as anteriores. Contudo, também o trabalho realizado nas várias etapas será expresso de acordo com a qualidade da representação efectuada. As disparidades que se podem observar devido a este processo têm, muitas vezes, uma magnitude superior ao das registadas aquando da alteração de algum aspecto nas etapas precedentes à representação, daí a elevada importância desta tarefa.

No entanto, apesar dos inúmeros trabalhos em avaliação de susceptibilidade a movimentos de vertente que procuram comparar diferentes modelos de análise, escalas de trabalho, número de variáveis, entre outros parâmetros que podem auxiliar na melhora de resultados de um modelo, só muito raramente se encontra uma referência ao método como foi executada a classificação de dados que originou determinado mapa. Com o aparecimento dos Sistemas de Informação Geográfica (SIG) e de *softwares* estatísticos, as divisões em classes são realizadas de uma forma rápida, quando não automática e displicente, não havendo qualquer interpretação crítica dos seus resultados.

Assim, através deste artigo pretende-se: (i) demonstrar como a etapa de classificação de dados pode influir no mapa de susceptibilidade; e (ii) indicar uma forma de avaliar a representação mais adequada em função dos objectivos da investigação a realizar, através da comparação de vários mapas de susceptibilidade a movimentos de vertente na bacia da Abadia (Torres Vedras).

2. ENQUADRAMENTO GERAL DA ÁREA DE ESTUDO E METODOLOGIA

A depressão da Abadia é uma pequena bacia hidrográfica (7 km²), situada cerca de 4 km a Este de Torres Vedras (Fig. 2).

Segundo Garcia (2002), a área em estudo corresponde estruturalmente a um anticlinal esventrado, de direcção NE-SW, estando o eixo ligeiramente deslocado para NW em relação ao centro da depressão. Talhada maioritariamente em margas e argilas (Complexo da Abadia), a depressão é dominada por cornijas calcárias (Complexo Corálico do Amaral), sendo ambas as formações de idade Kimeridgiana (Jurássico Superior).

A área em estudo, caracteriza-se, em termos morfológicos, por vertentes de perfil longitudinal côncavo, por vezes interrompido por rechãs, associadas a intercalações de arenitos no Complexo da Abadia. Estas vertentes apresentam uma dissimetria, sendo a vertente exposta a SE de maior comprimento e declive médio inferior. Contudo, devido ao contacto falhado entre o complexo da Abadia e os calcários do Amaral, o sector montante desta vertente apresenta dos declives mais acentuados de toda a área em estudo (superiores a 25^º). No conjunto, são os declives moderados (5^º-15^º) que predominam, sendo observados em cerca de 50 % da depressão.

A bacia (ordem 4 de acordo com a hierarquia de Strahler) drenada pelo rio Sangue, único curso de água permanente, apresenta uma densidade de drenagem de cerca 4 km/km².

O levantamento geomorfológico da área, baseado em interpretação de fotografias aéreas (1:15 000 e 1:26 000), ortofotomapas (1:10 000) e observações de campo apoiadas em cartografia de grande escala (1:2 000), realizadas entre Março de 2000 e Maio de 2001, permitiu identificar 105 movimentos de

vertente (94 000 m² de área instabilizada), dos quais 44 são deslizamentos rotacionais (Fig. 3), que representam 73,5 % da superfície instabilizada na bacia.

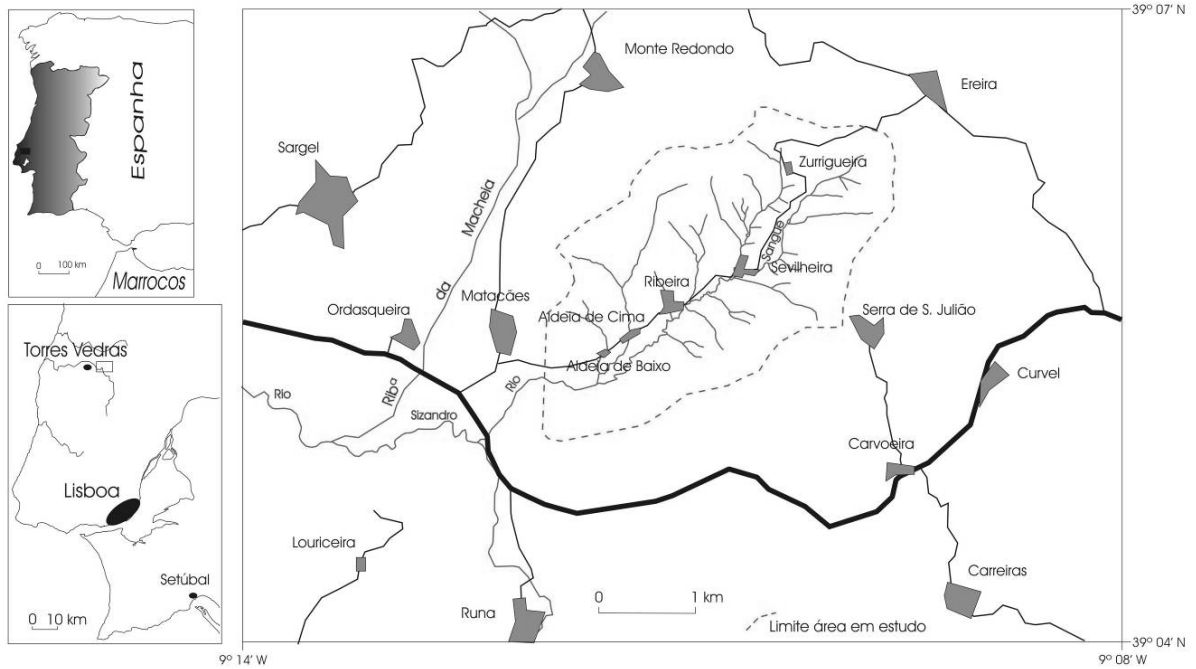


Figura 2 – Localização da área de estudo (Torres Vedras, Portugal).

Figure 2 – Location of the study area (Torres Vedras, Portugal).

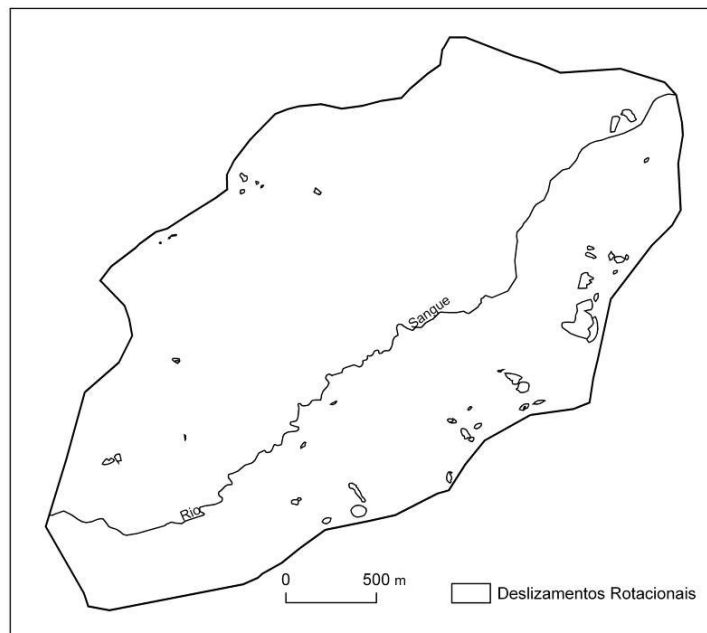


Figura 3 – Distribuição dos deslizamentos rotacionais na Bacia da Abadia.

Figure 3 – Rotational landslides distribution in Abadia Basin.

O Quadro 1 apresenta as variáveis consideradas no processo de avaliação de susceptibilidade aos movimentos de vertente rotacionais, tendo sido registados, através da sobreposição dos mapas de cada variável, para cada uma das 265 910 unidades de terreno (UT = pixel 5 x 5 m), quais as classes presentes.

Quadro 1

Factores condicionantes considerados na avaliação da susceptibilidade a deslizamentos rotacionais na Bacia da Abadia.

Table 1
Predispositional factors used for rotational landslides susceptibility assessment on the Abadia Basin.

Variáveis <i>Variables</i>	Classes <i>Classes</i>
Litologia <i>Litology</i>	UL1 – Margas e argilas
	UL2 – Arenitos e conglomerados
	UL3 – Calcários
	UL4 – Calcários com intercalações greso-margosas
	UL5 – Basaltos
	UL6 – Aluviões
Declive <i>Slope angle</i>	D1 – 0° - 2°
	D2 – 2° - 5°
	D3 – 5° - 10°
	D4 – 10° - 15°
	D5 – 15° - 20°
	D6 – 20° - 25°
	D7 – 25° - 30°
	D8 - > 30°
Perfil transversal da vertente <i>Transversal profile of slope</i>	P1 – Rectilíneo
	P2 – Côncavo
	P3 – Convexo
	P4 – Áreas planas
Morfoestrutura <i>Structure</i>	M1 – Vertente cataclinal
	M2 – Outras vertentes
Exposição <i>Slope aspect</i>	E1 – N
	E2 – NE
	E3 – E
	E4 – SE
	E5 – S
	E6 – SW
	E7 – W
	E8 – NW
	E9 – Sem exposição
Coberto vegetal e uso do solo <i>Land use and land cover</i>	CVUS1 – Inculto (herbáceo)
	CVUS2 – Inculto (arbóreo) e matagal
	CVUS3 – Cultivado (cereais)
	CVUS4 – Cultivado (vinha)
	CVUS5 – Socalcos (sem plantação ou plantação recente)
	CVUS6 – Vegetação porte arbóreo
	CVUS7 – Povoações e infra-estruturas antrópicas
Cursos de água <i>Streams</i>	R1 – Presença
	R2 – Ausência
Estradas / caminhos <i>Roads (municipality / unclassified)</i>	EC1 – Presença
	EC2 – Ausência

A avaliação da susceptibilidade à ocorrência de movimentos rotacionais, foi realizada utilizando o processo da análise discriminante. Este método estatístico multivariado pretende identificar inter-relações entre indivíduos de uma população e através destas explicar o comportamento de um determinado grupo, isto é, através das relações existentes entre os factores condicionantes “prever” se determinada unidade de terreno é susceptível de apresentar instabilidades de cariz rotacional ou não, minimizando os erros de classificação. Assim, registam-se para cada movimento os vários parâmetros que caracterizam os terrenos, as chamadas variáveis aleatórias multidimensionais (Santacana, 2001, p. 85), sendo estes que o modelo terá em conta para justificar os diferentes comportamentos e assim calcular a função discriminante. A função discriminante é obtida através da seguinte fórmula:

$$D = d_1V_1 + d_2V_2 + \dots + d_nV_n \quad (\text{Eq. 1})$$

sendo: D – valor discriminante da função; d_i – coeficientes de classificação estimados; V_i – variáveis independentes.

Da aplicação da técnica à Bacia da Abadia resulta a atribuição de um *score* indicativo da susceptibilidade para cada uma das 265 910 unidades que constituem a área em estudo. Os resultados obtidos através da aplicação da análise discriminante, efectuada através do processo *Forward Stepwise*, aos movimentos rotacionais da depressão da Abadia são resumidamente apresentados no Quadro 2.

Quadro 2

Matriz de classificação das unidades de terreno quanto à susceptibilidade a deslizamentos rotacionais na Bacia da Abadia, resultante da análise discriminante.

Table 2

Terrain units classification matrix resulting from rotational slides susceptibility assessment in the Abadia Basin, using discriminant analysis.

Grupo actual <i>Real group</i>	Grupo predito (%) <i>Predicted group (%)</i>	
	Estável <i>Stable</i>	Instável <i>Unstable</i>
Estável <i>Stable</i>	87,9	13,1
Instável <i>Unstable</i>	40,3	59,7

Casos correctamente classificados/Correctly classified cases: 87,7 %.

A percentagem de casos correctamente classificados é muito satisfatória (87,7 %). As unidades de terreno actualmente estáveis mas classificadas como instáveis (13,1 %) podem indicar locais onde futuramente ocorrerão deslizamentos, uma vez que estas apresentam características idênticas às actualmente instáveis (59,7 %). Se, de uma forma geral, as unidades estáveis são bem discriminadas (87,9 %) o mesmo não acontece com as unidades instáveis onde cerca de 40% são mal classificadas. No entanto, estes resultados seriam de esperar uma vez que o método apresenta algumas limitações, sobretudo quando não se teve o cuidado de só utilizar variáveis que não se interrelacionem e apresentem distribuições normais (Norusis, 1985 *in* van Asch & Mulder, 1991; Santacana, 2001; Garcia, 2002), não sendo a discussão destes resultados objecto deste trabalho.

Não obstante os resultados menos conseguidos na classificação automática de unidades de terreno, ao validarmos a fiabilidade das inter-relações entre variáveis utilizadas no modelo e sua eficácia preditiva,

através da curva de sucesso (Fig. 4), verifica-se que os resultados são bastante aceitáveis. Esta curva permite, ainda, estimar as probabilidades de ocorrência de futuros movimentos, por classe de susceptibilidade, facto indispensável para se utilizarem estes mapas no cálculo do risco (Fabri *et al*, 2002).

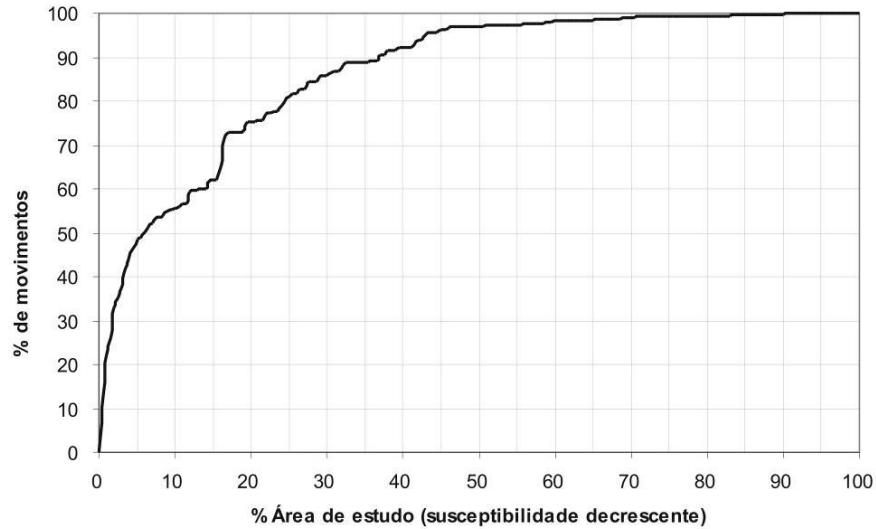


Figura 4 – Curva de sucesso do método da análise discriminante aplicado aos deslizamentos rotacionais na Bacia da Abadia.

Figure 4 – Success rate from discriminant analysis applied to rotational landslides of Abadia Basin.

Na construção destas curvas é necessário calcular dois rácios com as seguintes formulações:

$$R\acute{a}cio1 = \frac{Us}{Ut} * 100 \quad (\text{Eq. 2}) \quad R\acute{a}cio2 = \frac{Ums}{Um} * 100 \quad (\text{Eq. 3})$$

sendo que o rácio 1 relaciona o número de unidades de terreno seleccionadas (U_s) com o número total de unidades da área em estudo (U_t), i.e., a percentagem da área estudada, geralmente representada nas abcissas por ordem decrescente de susceptibilidade. O rácio 2 calcula a percentagem de movimentos validados, relacionando o número de unidades seleccionadas com movimentos (U_{ms}) e o número total de unidades instabilizadas na área estudada (U_m).

Assim, através da análise do gráfico pode-se constatar que com os 5 % das unidades de maior susceptibilidade se conseguem justificar quase 50 % dos movimentos. Por outro lado, se seleccionarmos a metade de unidades de terreno mais susceptíveis, consegue-se abranger mais de 97 % dos movimentos ocorridos, o que é na realidade um resultado bastante satisfatório.

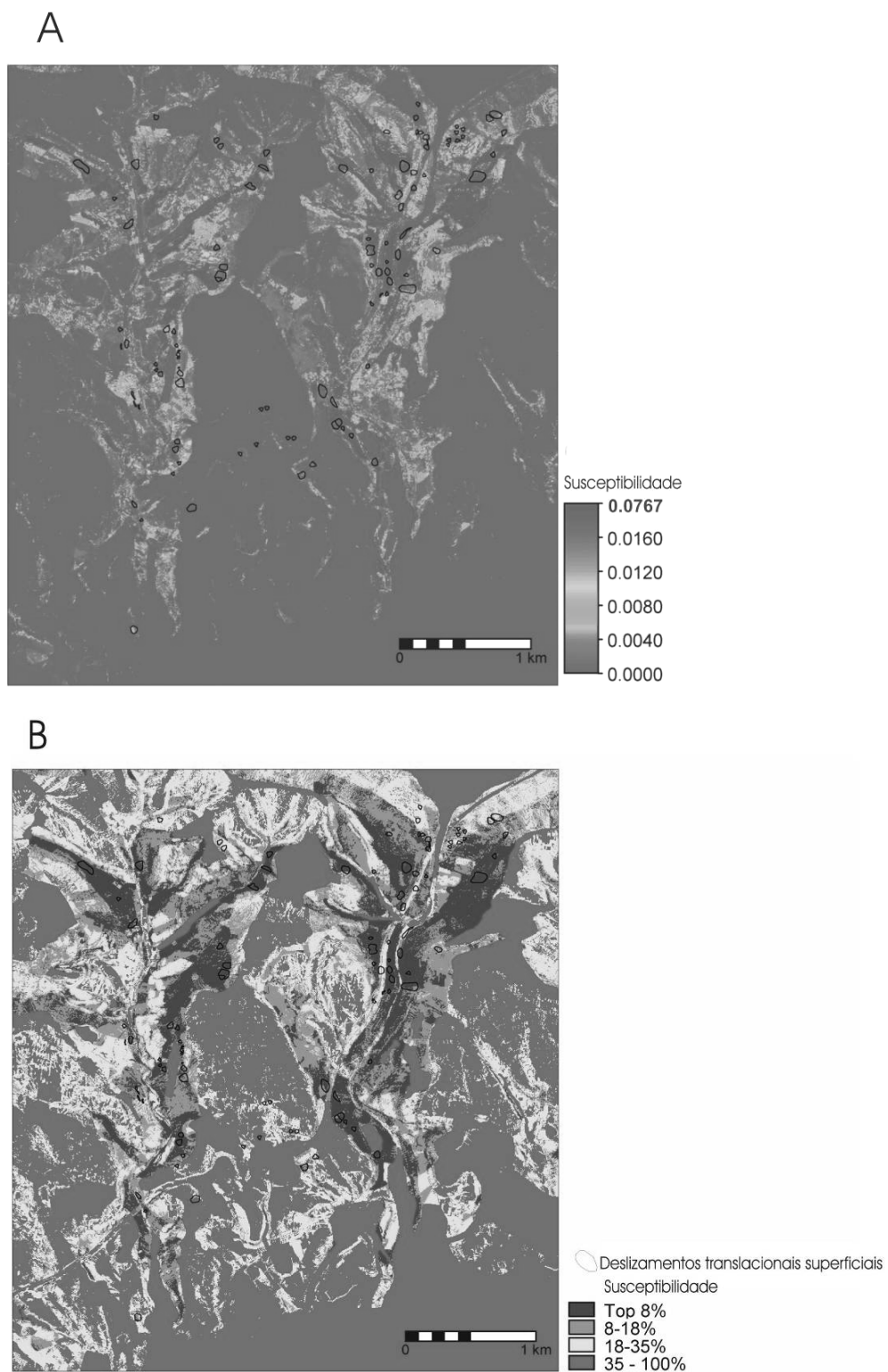


Figura 5 – Exemplos de mapas de susceptibilidade a deslizamentos. (A) Não classificado, (B) classificado (Adaptado de Zêzere et al, 2004).

Figure 5 – Shallow translational slides susceptibility maps examples. (A) Non classified map, (B) Classified map (Zêzere et al, 2004 adaptation).

3. MAPAS DE SUSCEPTIBILIDADE E O PROCESSO DE CLASSIFICAÇÃO DE DADOS

O processo de classificação de dados, vulgo divisão em classes, é sem dúvida uma etapa extremamente importante na elaboração de mapas de susceptibilidade, quer seja dos dados de base do modelo quer dos seus resultados. Como se pode observar na figura 5, a não classificação dos dados pode levar a uma interpretação difícil ou tornar inadaptado um trabalho de investigação rigoroso e coerente quando se pensa na sua aplicação prática. Na realidade o mapa classificado (B) apresenta uma simplificação e menor pormenor do que o mapa não classificado (A). Contudo, na maioria das situações práticas da sua aplicação, a perda de informação que se verifica devido à generalização e agrupamento de dados realizada no processo de divisão em classes, não é significativa tendo em conta o objectivo final do mapa. No exemplo estudado a seguir, utilizam-se sempre os mesmos dados de base, i.e. variáveis e modelo estatístico não sofrem qualquer modificação; o que muda é apenas o método de divisão em classes utilizado, procurando-se que o mapa de susceptibilidade final tenha 4 classes.

No Quadro 3 podem observar-se os limites das classes para a elaboração dos mapas de susceptibilidade a deslizamentos rotacionais, de acordo com 6 métodos de classificação de dados, formal e estatisticamente correctos, que se descrevem em seguida:

- A) Igual número de UT – método que divide a área em 4 partes iguais (quartis), apresentando cada classe o mesmo número de unidades de terreno;
- B) Igual amplitude – tendo em conta os valores máximo e mínimo dos *scores*, resultantes da análise discriminante, criam-se 4 classes com igual amplitude;
- C) Distribuição de todas as UT – limitação das classes baseada no diagrama de dispersão de todos os valores dos *scores* discriminantes;
- D) Distribuição das UT instáveis – método assente na dispersão dos *scores* das unidades que apresentam instabilidade;
- E) UT instáveis nos níveis elevados – adaptação do método do diagrama de dispersão, forçando que todas as unidades de terreno com instabilidade se encontrem nas duas classes de susceptibilidade mais elevada (forte e muito forte);
- F) Curva de sucesso – os limites de classe são demarcados com base na análise das rupturas de declive da curva de sucesso (Fig. 4) e dos valores que a originaram.

Os dois primeiros métodos são processos estritamente estatísticos, onde quem realiza o estudo não terá qualquer influência, para além da definição do número de classes, pelo que se podem considerar de subjectividade reduzida. No entanto, o mesmo não ocorre nos restantes métodos onde, quer seja baseado nas distribuições dos *scores* discriminantes das UT ou nos valores da taxa de sucesso, a influência de quem realiza esta operação nos resultados finais é elevada, pelo que a sua subjectividade é superior aos métodos puramente estatísticos.

Numa análise geral dos resultados facilmente se depreende que cada modelo tem valores limite de classes muito diferenciados, não obstante a manutenção do número das mesmas.

Analisando a figura 6, que representa os resultados da aplicação dos vários métodos de classificação de dados aos mapas de susceptibilidade a deslizamentos rotacionais na Bacia da Abadia, verifica-se que as diferenças são muito significativas. Na realidade, ao compararmos, por exemplo, o mapa em que todas as

unidades de terreno instáveis se encontram nas classes de susceptibilidade forte ou muito forte, observa-se visualmente nos mostra uma realidade completamente diferente da dos outros mapas.

Com base na figura 6 duas questões se colocam: “Qual é o mapa mais correcto?” e “Qual o melhor mapa?”.

Quadro 3

Limites de classes de susceptibilidade obtidos segundo diferentes métodos de classificação dos *scores* da análise discriminante aplicada aos deslizamentos rotacionais da Bacia da Abadia.

Table 3

Susceptibility classes limits obtain by different discriminant analysis scores classification to rotational landslides in the Abadia Basin.

	A	B	C
Classes de susceptibilidade <i>Susceptibility Classes</i>	Igual número de UT <i>Equal number of TU</i>	Igual amplitude <i>Equal range</i>	Distribuição de todas as UT <i>TU distribution</i>
Reduzida ou nula <i>Low or null</i>	≥ 0,6399	≥ 0,3845	≥ 1,0005
Moderada <i>Moderate</i>	0,3055 a 0,6398	-1,2464 a 0,3844	0,0005 a 1,0004
Forte <i>High</i>	-0,3574 a 0,3054	-2,8764 a -1,2465	-0,9994 a 0,0004
Muito forte <i>Very high</i>	≤ -0,3575	≤ -2,8765	≤ -0,9995
	D	E	F
Classes de susceptibilidade <i>Susceptibility classes</i>	Distribuição das UT Instáveis <i>Unstable TU distribution</i>	UT instáveis nos níveis elevados <i>Unstable TU in high susceptibility classes</i>	Curva de sucesso <i>Success rate curve</i>
Reduzida ou nula <i>Low or null</i>	≥ 0,0005	≥ 2,0005	≥ 0,1825
Moderada <i>Moderate</i>	-0,9994 a 0,0004	0,9995 a 2,0004	-0,5514 a 0,1824
Forte <i>High</i>	-2,9994 a -0,9995	-0,9994 a 0,9994	-2,4224 a -0,5515
Muito forte <i>Very high</i>	≤ -2,9995	≤ -0,9995	≤ -2,4225

UT – Unidades de Terreno/ TU – Terrain Units.

A resposta à primeira questão, e tendo em conta exclusivamente as técnicas utilizadas para a elaboração dos mapas, seria “Todos”. Na verdade, não existe um mapa mais “correcto” que os outros, desde que as técnicas sejam bem aplicadas; pode é, neste caso, haver métodos que serão mais adaptados à realidade que estudamos. Pode ser desaconselhável, ou menos lógico, utilizar métodos puramente estatísticos neste tipo de mapas, uma vez que o critério de divisão em classes não tem qualquer relação com os tipos de variáveis que se analisam, mas o procedimento não está errado. As metodologias que têm em conta a existência de movimentos são, à partida, aquelas que representam uma melhor imagem da realidade, uma vez que se baseiam em factos reais e não apenas em análises de frequência ou de amplitude. Incluem-se nestes métodos: a classificação de dados baseada na distribuição dos *scores* das unidades de terreno instabilizadas; a divisão, mais conservadora, que engloba nas duas classes mais susceptíveis todas as unidades com movimentos; e a interpretação das taxas de sucesso.

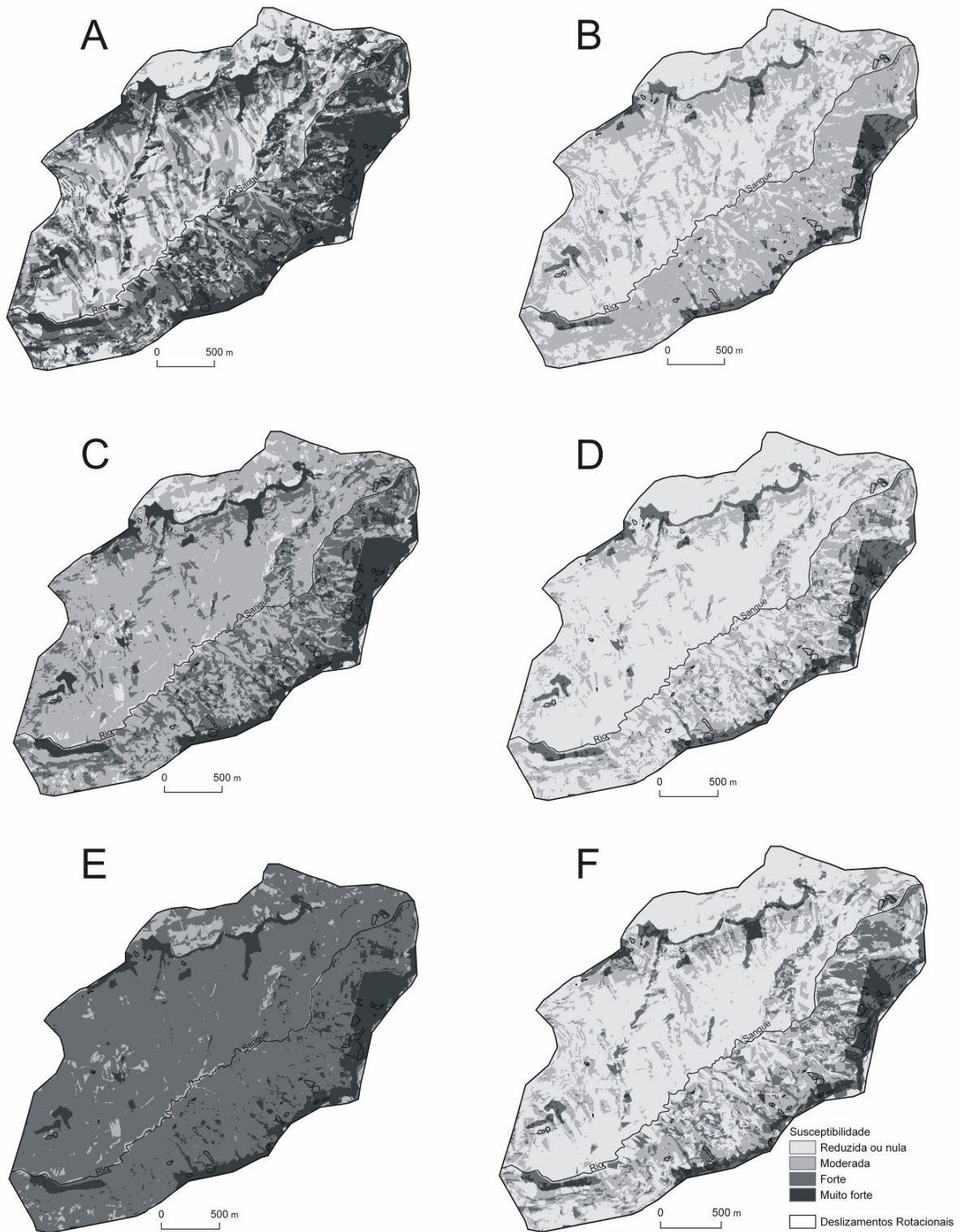


Figura 6 – Mapas de susceptibilidade a deslizamentos rotacionais na Bacia da Abadia utilizando diferentes métodos de classificação de dados. **A** – Igual número de TU; **B** – Igual amplitude; **C** – Distribuição de todas as UT; **D** – Distribuição das UT instáveis; **E** – UT instáveis nos níveis elevados; **F** – Curva de Sucesso.

Figure 6 – Rotational slides susceptibility maps in the Abadia Basin using different methods of data classification. **A** – Equal number of TU; **B** – Equal range; **C** – TU distribution; **D** – Unstable TU distribution; **E** – Unstable TU in High susceptibility classes; **F** – Success rate curve.

No que diz respeito à segunda questão, a resposta é mais complicada e dependerá sempre do objectivo que se pretende atingir. Contudo, o mapa que tem todos os movimentos nas áreas de susceptibilidade mais elevada não é obrigatoriamente o melhor, pelo contrário, principalmente se essa condição acarretar a existência de um elevado número de “falsos positivos” (áreas sem movimentos classificadas como instáveis), como acontece no mapa E (Fig. 6). O mapa ideal é aquele que consegue ter o maior número de movimentos nas classes mais susceptíveis, mas tendo estas a menor área possível, aumentando deste modo a respectiva capacidade preditiva. Uma análise de “custo-benefício” do alargamento da área de uma classe torna-se imprescindível aquando da aplicação destes mapas ao planeamento e ordenamento do território.

Tendo em conta esta premissa e a distribuição dos movimentos, em termos visuais, facilmente se conclui o exagero da área ocupada pelas classes moderada, forte e muito forte do mapa C, com a quase inexistência de áreas com susceptibilidade reduzida ou nula. Este resultado deve-se à distribuição bimodal dos *scores* das unidades instabilizadas e à adopção de um critério onde a probabilidade de errar é quase nula, revelando-se demasiado conservador (Garcia, 2002, p. 116).

Contudo, só com um conhecimento de campo muito pormenorizado, e mesmo assim por vezes insuficiente, se poderiam avaliar de forma clara os restantes mapas, uma vez que todos apresentam problemas e virtudes.

A tabulação de dados e a utilização de índices que avaliem a capacidade preditiva das classes obtidas auxilia na escolha do mapa mais adequado aos objectivos, com menor subjectividade. No Quadro 4 apresentam-se os resultados de um índice que permite comparar a capacidade “preditiva” de cada classe (Eq. 4):

$$I = \log \left(\frac{Pd/Pdt}{Ac/At} \right) \quad (\text{Eq. 4})$$

sendo: Pd – número de pixels deslizados na classe; Pdt – total de pixels deslizados na área em estudo; Ac – área da classe (em número de pixels); At – área total em estudo (em pixels).

O índice de capacidade preditiva das classes (Eq. 4), resulta da ponderação entre a importância que a classe tem na explicação dos movimentos (Pd/Pdt) e a importância da classe na área total (Ac/At), sendo que tanto melhor é o índice quanto mais elevado o seu valor. A utilização do logaritmo natural permite uma normalização dos dados e, por outro lado, possibilita uma comparação dos resultados de cada classe com a densidade média de movimentos da área, i.e., classes com sinal positivo indicam maior justificação do que a média e classes com sinal negativo valores inferiores à média.

Assim, da análise do quadro 4 podemos mais uma vez concluir que não é a classe muito forte que tem mais pixels deslizados (2 261, no método “A”) aquele que apresenta uma melhor capacidade preditiva (0,51) uma vez que a área dessa classe é muito elevada. Pelo contrário, é o método que apresenta menor superfície instabilizada na classe mais susceptível (Método “F” - 908 pixels) que tem o melhor índice de predição (1,22) uma vez que a área dessa classe é a mais diminuta. Ao avaliarmos as duas classes de susceptibilidade mais elevada verificamos que os maiores valores são obtidos através da curva de sucesso (Método “F”: MF=1,22; F=0,43), seguido pela distribuição das unidades de terreno instáveis (Método “D”: MF=1,14; F=0,37) e classificação por igual amplitude (Método “B”: MF=1,10; F=0,34). Se os

dois primeiros métodos, embora de maior subjectividade, não surpreendem, tendo em conta que se baseiam na distribuição das áreas deslizadas, o processo de classificação de igual amplitude é puramente estatístico, pelo que o bom resultado do método pode dever-se a casualidade estatística necessitando, por isso, de validação com outros estudos comparativos.

Quadro 4
Índice de capacidade preditiva das classes de susceptibilidade a deslizamentos rotacionais da Bacia da Abadia.

Table 4
Predictive index of susceptibility classes to rotational landslides in the Abadia Basin.

Classes de susceptibilidade <i>Susceptibility classes</i>	A Igual número de UT <i>Equal number of TU</i>			B Igual amplitude <i>Equal range</i>			C Distribuição de todas as UT <i>TU distribution</i>		
	Pd	Ac*	I	Pd	Ac	I	Pd	Ac	I
Reduzida ou nula <i>Low or null</i>	22	66 544	-1,50	74	121 245	-1,23	0	16 893	
Moderada <i>Moderate</i>	58	66 752	-1,08	1 157	117 496	-0,02	304	153 820	-0,72
Forte <i>High</i>	422	65 752	-0,21	421	18 694	0,34	809	61 918	0,10
Muito forte <i>Very high</i>	2 261	66 862	0,51	1 111	8 475	1,10	1 650	33 279	0,68
Área abaixo da curva (AAC) <i>Area under the curve</i>		0,820			0,827			0,821	

*A área de cada classe não é exactamente igual devido ao valor do limite de classe ser registado em várias unidades de terreno que não devem ser separadas tendo as mesmas características.

Classes de susceptibilidade <i>Susceptibility Classes</i>	D Distribuição das UT Instáveis <i>Unstable TU distribution</i>			E UT instáveis nos níveis elevados <i>Unstable TU in High susceptibility classes</i>			F Curva de sucesso <i>Success rate curve</i>		
	Pd	Ac	I	Pd	Ac	I	Pd	Ac	I
Reduzida ou nula <i>Low or null</i>	304	170 713	-0,77	0	7		86	142 263	-1,24
Moderada <i>Moderate</i>	809	61 918	0,10	0	16 895		665	78 443	-0,09
Forte <i>High</i>	630	26 091	0,37	1 113	215 729	-0,30	1 104	39 887	0,43
Muito forte <i>Very high</i>	1 020	7 188	1,14	1 650	33 279	0,68	908	5 317	1,22
Área abaixo da curva (AAC) <i>Area under the curve</i>		0,832			0,749			0,860	

UT – Unidades de Terreno/ TU – Terrain Units; Pd – N.º pixels deslizados; Ac – N.º pixels da classe; I – Índice.

No entanto, estes índices apenas permitem a comparação entre classes. A avaliação da qualidade preditiva global de cada um dos métodos utilizados foi efectuada recorrendo à construção das respectivas curvas preditivas (Fig. 7) e ao cálculo das “Áreas Abaixo da Curva (AAC)” correspondentes (Gorservski et al, 2000; Bi & Bennett, 2003), através da seguinte fórmula (Eq. 5):

$$AAC = \sum_{i=1}^n \left[(L_{s_i} - L_i) * \frac{a_i + b_i}{2} \right] \quad (\text{Eq. 5})$$

sendo: $(Ls_i - L_i)$ – amplitude da classe; a_i – valor da ordenada correspondente a L_i ; b_i – valor da ordenada correspondente a Ls_i .

Os resultados da aplicação da equação (Eq. 5) estão expressos no Quadro 4. As AAC podem, teoricamente, variar entre 0 (mínima capacidade preditiva), e 1 (máxima capacidade preditiva). O valor de 0,5 representa o mínimo exigível a um modelo; isto é, por exemplo, com 70% da área estudada prever 70% dos movimentos. Porém, deve ter-se em conta que o valor total obtido através da Eq. 5, não dispensa a respectiva interpretação dos valores de cada classe.

Da análise da Figura 7 e dos dados do Quadro 4, verifica-se que o método cuja divisão em classes se baseia na curva de sucesso é o que apresenta o melhor resultado (AAC = 0,860). Por sua vez, o método em que todos os movimentos são classificados nas classes de maior susceptibilidade (E) é o que apresenta o pior resultado (AAC = 0,749). Assim, comprova-se que o mapa que apresenta melhor relação entre a dimensão de áreas susceptíveis e o número de movimentos explicados é o baseado na curva de sucesso. Contudo, convém ter presente que a qualidade preditiva de qualquer modelo de susceptibilidade depende, acima de tudo, da qualidade dos dados usados na sua construção, nomeadamente a fiabilidade do inventário dos movimentos de vertente e o significado das variáveis independentes utilizadas enquanto factores condicionantes da instabilidade geomorfológica.

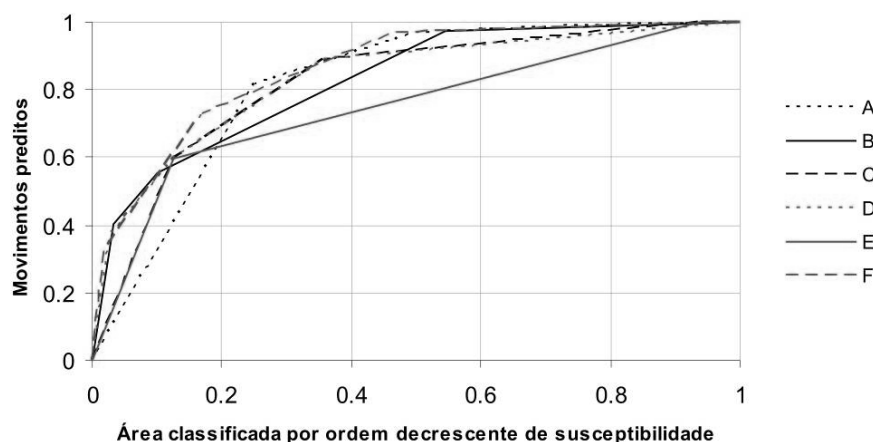


Figura 7 – Capacidade preditiva dos métodos de classificação de dados aplicados aos deslizamentos rotacionais na Bacia da Abadia. **A** – Igual número de UT; **B** – Igual amplitude; **C** – Distribuição de todas as UT; **D** - Distribuição das UT instáveis; **E** – UT instáveis nos níveis elevados; **F** – Curva de Sucesso.

Figure 7 – Prediction variation of classification data methods applied to rotational landslides of Abadia Basin. **A** – Equal number of TU; **B** – Equal range; **C** – TU distribution; **D** – Unstable TU distribution; **E** – Unstable TU in High susceptibility classes; **F** – Success rate curve.

4. CONCLUSÃO

Os resultados obtidos neste trabalho comprovam que o processo de classificação de dados em classes é uma etapa muito importante e indispensável para a expressão cartográfica da avaliação de susceptibilidade, influenciando fortemente o produto final. De facto, aplicando o mesmo modelo preditivo poder-se-ão elaborar mapas completamente diferentes. Para tal, basta modificar o método de divisão em classes utilizado, sendo que qualquer deles é formalmente correcto. O melhor mapa de susceptibilidade não é o que apresenta maiores áreas nos níveis de susceptibilidade elevada, mas aquele que apresenta

uma melhor relação entre o maior número de movimentos preditos e a menor percentagem de área classificada como mais susceptível, contribuindo este facto para o aumento da respectiva capacidade preditiva. Para se identificar qual o melhor método é necessário avaliar esta relação para cada um dos mapas elaborados, com base em processos distintos de classificação de dados. A utilização de índices auxilia este processo mas não substitui uma análise crítica dos resultados, tendo em conta os objectivos a atingir e o indispensável conhecimento de campo da área em estudo. No trabalho apresentado, os valores de susceptibilidade são obtidos através da análise discriminante aplicada aos deslizamentos rotacionais da Bacia da Abadia. O processo de divisão em classes que origina mapas com melhores resultados é o método baseado na taxa de sucesso. O facto de ser um modo de classificação de dados mais subjectivo, quando comparado com os exclusivamente estatísticos, não lhe retira rigor ou correcção. Deste modo, fica provado que esta etapa de trabalho deve ser tida em conta e não realizada de forma displicente, uma vez que da sua correcta execução dependerá a boa interpretação dos resultados, podendo uma má opção na escolha da metodologia de divisão em classes fazer com que todo o trabalho realizado tenha sido em vão. O uso adequado dos mapas e a sua utilidade são fortemente condicionados pelo procedimento de classificação de dados adoptado.

BIBLIOGRAFIA

- Bi, J. & Bennett, K. P. (2003) – Regression Error Characteristic Curves. *Proceedings of the Twentieth International Conference on Machine Learning (ICML-2003)*, Washington DC, 8p.
- Carrara, A. L. (1993) – Uncertainty in evaluating landslide hazard and risk. In Nemeč, J.; Nigg, J. M.; Siccardi, F. (Eds.) – *Predictions and Perception of Natural Hazards*, Kluwer Academic Publishers, Dordrecht, p. 101-109.
- Fabbri, A. G.; Chung, C. F.; Napolitano, P.; Remondo, J.; Zêzere, J. L. (2002) – Prediction rate functions on landslide susceptibility applied in the Iberian Peninsula. *Risk Analyses 2002, Conference on computer simulation in risk analyses and hazard mitigation*.
- Garcia, R. A. C. (2002) – *Avaliação do Risco de Movimentos de Vertente na Depressão da Abadia (Torres Vedras)*. Tese de mestrado em Geografia Física e Ambiente, Universidade de Lisboa, 140p. + anexo cartográfico.
- Gorsevski, P. V.; Gessler, P.; Foltz, R. B. (2000) – Spatial Prediction of Landslide Hazard Using Logistic Regression and GIS. *4th International Conference on Integrating GIS and Environmental Modeling (GIS/EM4): Problems, Prospects and Research Needs*. Banff, Alberta, Canada
- Santacana, N. (2001) – *Análisis de la Susceptibilidad del Terreno a la Formación de Deslizamientos Superficiales y Grandes Deslizamientos Mediante el Uso de Sistemas de Información Geográfica. Aplicación a la Cuenca Alta del Río Llobregat*. Tese de doutoramento em Engenharia del Terreny i Cartogràfica, Universitat Politècnica de Catalunya, Barcelona, 399p.
- van Asch, T. W. J. & Mulder, H. F. H. M. (1991) – Statistical, geotechnical and hydrological approaches in landslide hazard assessment of mass movements. In UNESCO-ITC *Project on Mountain Hazard Mapping in Andean Environment using Geographical Information Systems*, Expert Workshop, Bogotá, 31p.
- Zêzere, J. L.; Reis, E.; Garcia, R.; Oliveira, S.; Rodrigues, M. L.; Vieira, G.; Ferreira, A. B. (2004) – Integration of spatial and temporal data for the definition of different landslide hazard scenarios in the area north of Lisbon (Portugal). *Natural Hazards and Earth System Sciences*, vol. 4, European Geosciences Union, p. 133-146.

AGRADECIMENTOS

Os autores agradecem ao professor Lúcio Cunha a revisão do artigo, cujos comentários e sugestões conduziram à melhoria do artigo.