

# O geo-processamento da informação e seu tratamento estatístico no âmbito do projecto Montado

Paulo M. Pereira & Miguel Pereira

## Sumário

O Montado é um sistema agro-silvo-pastoril que constitui simultaneamente um ecossistema natural de grande diversidade e uma fonte de recursos e riqueza para o homem. A sua compreensão requer cada vez mais uma abordagem multidisciplinar capaz de abarcar todas as dimensões deste complexo ecossistema.

A informação geográfica é cada vez mais um utensílio indispensável para a planificação e gestão do território. A correcta estruturação dos dados representa um desafio crescente nomeadamente no domínio do conhecimento georreferenciado. A combinação de Sistemas de Informação Geográfica (SIG) com ferramentas estatísticas adequadas permite tratar e visualizar grandes quantidades de informação, filtrando-a e simplificando-a para a sua posterior utilização; posteriormente, a aquisição de informação, o seu processamento e tratamento estatístico permite gerar nova informação, crucial para a boa gestão do Montado.

A metodologia desenvolvida para o Montado demonstrou ser eficaz para obter uma classificação do Alentejo baseada em 6 grupos de variáveis (climatologia, usos de solo, tipo de floresta, economia, gado e litologia), permitindo a estratificação da escolha das herdades e garantindo assim que estas fossem representativas dos diferentes tipos de Montado existentes no Alentejo. A ferramenta de tratamento e visualização de dados (SIG e estatística) permite por sua vez extrapolar os dados recolhidos nas herdades para toda a região Alentejana.

## Introdução

### ***O Montado***

O Montado é representativo de um sistema de exploração extensiva que concilia interesses humanos com a conservação da natureza. O Montado cobre hoje aproximadamente 6 milhões de hectares, sendo um ecossistema agro-silvo-pastoril com um coberto arbóreo pouco denso (60 a 100 árvores por hectare) dominado por Sobreiro (*Quercus suber*) ou/e Azinheira (*Quercus rotundifolia*), com pastagens e campos agrícolas no sub-coberto, normalmente com um sistema de rotação que alterna folhas em pousio e outras exploradas agricolamente (Pinto-Correia 1993, Lourenço et al. 1998). A cobertura arbórea não é uniforme já que resulta na maior parte dos casos de regeneração natural, o que implica ter havido uma adaptação ecológica da densidade de Sobreiro e Azinheira às condições climáticas locais (Joffre et al. 1999).

A sustentabilidade do montado parece ser um atributo deste sistema semi-natural (Pinto-Correia 1993, Herzog 1998). No entanto, é ainda difícil definir e balizar critérios de sustentabilidade na agricultura e na silvicultura. Uma das medidas mais populares de sustentabilidade, a biodiversidade, está mal quantificada nos montados, e apenas temos

alguns conhecimentos relativamente ao impacto das acções humanas num determinado grupo biológico ou espécie (Santini, F. & Angulo, A. 2001). Apesar de ser um sistema muito diverso (De Miguel 1999, Carrión et al. 2000), no Mediterrâneo, a compreensão da dinâmica da biodiversidade e da quantificação das suas componentes só é possível tendo em conta o factor antropogénico (Blondel & Aronson 1999).

A escolha do tipo de produção no montado torna-se assim um factor chave para a sua gestão (Coelho 1996), estando os sistemas extensivos de exploração associados a altos níveis de diversidade, como por exemplo acontece com a riqueza de avifauna (Woodhouse et al. 2005). Por sua vez, a gestão da terra é o factor determinante para a sua riqueza em biodiversidade (Onipchenko & Semenova 1995, Leiva et al. 1997). Além disso, a aproximação multidisciplinar dos ecossistemas, juntando os factores sócio-económicos, agrícolas, biológicos e ambientais tem-se revelado a forma mais frutuosa de abordar esta complexa e interessante problemática (Balmford & Bond 2005).

## **Sistemas de Informação**

A integração de diferentes tipos de dados é cada vez mais uma realidade mas também um crescente desafio (Gahegan et al. 2000). Embora clássica, a questão: *O que existe em tal espaço?* é fundamental para uma gestão territorial adequada, e para uma acção que se espera em harmonia com os diferentes agentes do espaço. Atribuir a um espaço certos dados geradores de informação implica desde logo uma capacidade de abstracção, mas também o processamento adequado dos dados que permitem a geração de nova informação georreferenciada. A utilização de modernos meios tecnológicos afigura-se indispensável no tratamento de volumes de dados cada vez mais vastos e de diferentes origens. O advento dos Sistemas de Informação (SI) tem possibilitado o tratamento e gestão da informação estruturada, mas também e não menos importante, um contacto, mesmo que virtual, próximo da realidade.

Os Sistemas de Informação Geográfica (SIG) apareceram da necessidade de analisar e quantificar a crescente quantidade de informação cartográfica em formato digital que acompanhou de perto a revolução informática (Nyerges 1993). Os SIG são por definição ferramentas excelentes para manipular dados espacialmente referenciados, descrevendo essencialmente relações espaciais e permitindo sobrepor diferentes capas de informação geográfica; neste contexto, assistimos hoje a um alargamento do seu espectro de aplicabilidade, nomeadamente no que se refere à modelação ecológica e testes de hipóteses (Haslett 1990, Walker 1990, Goodchild et al. 1993).

Os SIG são a ferramenta ideal para armazenar, manipular e processar bases de dados que tenham uma componente espacial. No caso das bases de dados complexas, os SIG têm a capacidade extremamente útil de compatibilizar dados de natureza muito diversa, como dados florísticos, rede viária, climatologia, detecção remota, geologia, sócio-económicos ou tipos de uso do solo. É através da sobreposição geográfica de toda esta informação, (possível e facilitada pela construção de um SIG integrador), que podemos testar hipóteses e modelar para todo o território através de processos de extrapolação baseados em estatística multivariada.

## **Estatística aplicada**

Métodos estatísticos de descrição (ordenação), classificação (*K-means*) e extrapolação (Árvores de classificação e regressão) de objectos permitem a análise de matrizes complexas de dados e sua utilização para gerar novas ferramentas de gestão do território.

As modernas técnicas de ordenação vieram permitir a análise destas matrizes, sendo uma técnica que se ajusta com mais rigor à natureza dos dados georreferenciados (Legendre 1998); o espaço multivariado definido pela matriz de  $n$  variáveis  $\times$   $n$  objectos (os objectos podem ser as freguesias e as variáveis os descritores destas mesmas freguesias) permite uma melhor aproximação à definição empírica multidimensional do montado, sendo assim possível quantificar a sua natureza multidisciplinar.

A base computacional da análise multivariada é a álgebra de matrizes, que permite analisar e manipular bases de dados multidimensionais; o algoritmo que reduz os dados multidimensionais a um número de dimensões facilmente manipulável é o algoritmo de *Householder*. Esta transformação simplifica a complexidade inicial, reduzindo aos primeiros eixos uma parte importante da variância total da matriz inicial. Com poucos eixos que explicam muita variância é agora possível visualizar freguesias e variáveis num espaço reduzido; quando este espaço se limita a duas dimensões (que poderá ser qualquer uma das combinações possíveis dos quatro primeiros eixos), a resultante gráfica chama-se “biplot” (Ter Braak & Looman 1994).

Ao usar-se um método de ordenação, espera-se que a distância entre as objectos projectados no biplot sejam fiéis descritores das dissimilaridades entre freguesias (Prentice 1977). Estes métodos de ordenação que se limitam à exploração dos dados enquadram-se na denominada análise exploratória, e compreende métodos como por exemplo a análise de componentes principais (*ACP*) a análise de correspondência (*AC*) e a análise de coordenadas principais (*ACoP*). Estes métodos distinguem-se essencialmente pelo tipo de distância utilizada para descrever as posições relativas dos objectos a ordenar: distância euclidiana (*ACP*) distância do  $\chi^2$  (*AC*) ou uma matriz de distâncias, onde o tipo distância é decidido pelo investigador (*ACoP*) (Legendre 1998).

A classificação não hierárquica permite agrupar objectos com base nos seus atributos, sendo o utilizador a definir à priori o número de agrupamentos que quer efectuar. O processo recorre ao algoritmo de *K-means* (Legendre 1998), que encontra a solução minimizando a variância dentro dos grupos e maximizando a variância entre grupos.

A metodologia das árvores de regressão e classificação (Breiman et al, 1984) é apropriada quando se tem uma grande quantidade de variáveis explicativas e quando se espera que haja relações complexas entre as variáveis dependentes e independentes. Nas árvores de regressão, combinações diferentes de variáveis podem dar origem ao mesmo resultado, o que seria impossível utilizando métodos lineares. Nestes casos, as árvores de regressão reflectem melhor a realidade, e têm uma interpretação fácil e intuitiva (Clark & Pregibon, 1990). Nesta metodologia estatística, a variável dependente é dividida em dois grupos diferentes num processo semelhante ao utilizado pelo algoritmos de *K-means*, sendo seleccionada a variável que melhor explica a divisão obtida. Este processo é repetido sucessivamente, até chegar aos ramos finais da árvores, que no exemplo aqui dado reflectem o estado de sanidade do montado de sobre. Cada classe de sanidade corresponderá a uma combinação única de variáveis, obtida seguindo o “ramo” final da classe até à divisão original da árvore.

## ***O projecto Montado***

A combinação de um SIG integrador com as metodologias estatísticas apresentadas permitiu tratar a natureza multidisciplinar do montado em duas fases essenciais do projecto: na estratificação da amostragem conducente à escolhas das 31 herdades estudadas por todos os grupos participantes e na extrapolação dos dados resultantes do levantamento feito pelas diferentes equipas nestas herdades. A eficácia desta extrapolação, ainda que imperfeita em consequência do reduzido número de herdades estudadas, poderá ser uma grande mais-valia para a futura gestão do território no Alentejo, já que permitirá ter informação relativamente barata num curto espaço de tempo, o que se adequa bastante às necessidades gestão e planificação desta região. A integração dos dados da Extremadura (Espanha) nos modelos permitirá sem dúvida melhorar a sua performance. A estruturação de todos os dados em função da freguesia revelou-se uma solução adequada para a combinação de dados de natureza muito diversa (Ex. dados provenientes do INE, do inventário florestal e dados climatológicos) o tratamento estatístico e com mapas resultantes facilmente interpretáveis pelos decisores e administradores do território.

## **Metodologia**

### ***Sistema de Informação Geográfica***

O desafio que se nos colocou foi a gestão de dados geográficos para o projecto INTERREG/Montado. Como objectivo do Projecto estava a identificação de factores e variáveis de condicionamento da biodiversidade e produtividade dos montados, recorrendo à centralização de dados em Sistema de Informação Geográfica (SIG) e a modelos estatísticos de descrição (ordenação), classificação (K-means) e extrapolação (Árvores de classificação e regressão). No SIG foram integrados todos os dados recolhidos durante o Projecto e será a partir dele que é gerada a informação que alimenta o Sistema de Apoio à Decisão (SAD) final. O modelo SIG para o projecto foi desenvolvido na [Unidade de Macroecologia & Conservação](#) (UMC). O trabalho de campo decorreu nas herdades seleccionadas para amostragem localizadas em montados da região do Alentejo. O modelo de dados permitiu a caracterização exaustiva das amostragens de campo e sua integração no universo de dados existentes na área de estudo. As amostragens realizadas foram fonte de um vasto conjunto de dados relativos às diferentes espécies de fauna e flora. O modelo de SIG desenvolvido teve como missão permitir análises preliminares dos dados, gerar novas variáveis, facilitar a interpretação do tratamento estatístico e contribuir para a realização de outputs de difusão e disseminação de resultados. O SIG é a estrutura base de toda a geo-informação editada e manipulada ao longo do projecto INTERREG/Montado. Será também esta a estrutura que permitirá que o output final do Projecto possa continuar a ser actualizado mesmo após o fim do dito. Por tudo isto, o planeamento correcto e antecipado do SIG adquiriu uma importância extrema.

O modelo de dados para o SIG teve por base o desenvolvimento de um projecto ArcView® e ArcGIS®, com uma estrutura de directorias dedicadas às diferentes necessidades temáticas e suas especificidades (inputs/outputs). Até ao presente o

trabalho de manipulação de informação digital subdivide-se em duas categorias principais:

a) Selecção e sistematização de informação existente, tendo sido seleccionados e tratados os seguintes temas de enquadramento:

- (1) Limites administrativos (NUT, Concelhos e Freguesias)
- (2) Limite das cartas topográficas 1: 25 000 IgeoE

Os temas administrativos preparados e editados referem-se a toda a dimensão administrativa em vigor e correspondente à dimensão espacial utilizada pelo EUROSTAT.

b) Edição e integração de informação base de domínio público, categoria de informação que se estruturou em três componentes principais:

- (1) Informação temática: O trabalho de estruturação do SIG desenvolvido inclui um vasto conjunto de temas editados para a dimensão espacial do projecto e que estão agrupados segundo os seguintes conjuntos - Morfológicos, Climáticos, Bioclimáticos, Usos do solo e de Inventário Local.
- (2) Modelo de dados alfanuméricos (MSAccess®): o modelo de dados desenvolvido inclui os dados referentes às análises e levantamentos dos grupos para os quais já decorreram trabalhos de campo, Inventário Florestal e Dados estatísticos do INE (à freguesia). A selecção dos dados de espécies permite visualização espacial por via de uma ligação ODBC ao projecto ArcView.
- (3) Metadados: Segundo FGDC CSDGM, com catalogação dos diferentes temas de projecto.

O trabalho executado está em conformidade com as exigências cartográficas da proposta inicial, que compatibiliza os sistemas de referência ibéricos. Seguidamente apresentamos uma breve descrição metodológica dos principais temas gerados para o projecto montado (os dados do INE, entraram directamente nas folhas de cálculo, já que estavam referenciados à freguesia):

### **Morfologia**

Na base da informação geomorfológica está o Modelo Digital de Terreno (MDT), essencial para a derivação de dados e informação associada, tais como declives, exposição ou nos processos de derivação em modelos que estejam relacionados com a topografia. O MDT do Alentejo - 1Km foi editado tendo como fonte o GTOPO30 e a rede hidrográfica regional, um produto de distribuição livre, da responsabilidade da USGS. GTOPO30 é um MDT global com um intervalo horizontal em grid de 30 segundos (aproximadamente 1km). GTOPO30 foi derivado de um conjunto de fontes de dados topográficos em formato vectorial e raster.

### **Modelação climática**

As variáveis climáticas são centrais para a caracterização ambiental da área de estudo. No caso da *precipitação*, o SNIG (Sistema Nacional de Informação Geográfica) disponibiliza já Mapas Mensais de Distribuição da Precipitação para todo o Portugal Continental. No caso da *temperatura*, porém, existe apenas disponível o Atlas do

Ambiente (Direcção-Geral do Ambiente) no qual foi considerada apenas a temperatura anual por classes discretas. Para o Projecto Montado/Dehesa, contudo, foram necessárias variáveis climatológicas contínuas.

Para cada uma das variáveis interpoladas (temperatura média, máxima e mínima, evapotranspiração, etc) foram analisadas várias combinações de variáveis acessórias, tendo em conta a bibliografia existente. O processo de modelação da *temperatura*, no qual se interpolaram os dados em falta compatibilizando os dados discretos referidos com a variação de outras variáveis ambientais relacionadas com a temperatura, essencialmente variáveis derivadas do Modelo Digital de Terreno e da cartografia de localização.

Foram explorados vários métodos de interpolação de dados: co-kriging esférico, circular, exponencial, etc. A escolha final do método e variáveis a usar para entrada em cada um dos modelos teve em conta o menor erro estatístico (por validação cruzada) dos modelos gerados. Como exemplo, para a modelação das Temperaturas médias, máximas e mínimas mensais foram usadas como co-variáveis: o MDT, o “Custo ao Mar” (função de “Custo” usando como peso o MDT) e a Precipitação média mensal. O método de co-kriging usado foi o método exponencial.

Informação utilizada:

- (1) Dados climáticos: Informação alfanumérica de 98 estações climáticas de Portugal Continental, de um período médio de 30 anos, resumidos ao mês.
- (2) Modelo Digital de Terreno e derivados: Abrangendo Portugal Continental, com uma resolução espacial de 1Km.
- (3) Mapas Mensais de Distribuição da Precipitação em Portugal Continental: Abrangendo Portugal Continental, com uma resolução espacial de 1Km.
- (4) Cartografia de localização: Limites administrativos de Portugal Continental e limite da Península Ibérica.

### **Modelação Bioclimática**

O tema de aridez foi calculado com base na normal climática de 30 anos (1961-1990). Foi também objecto de tratamento um conjunto de variáveis derivadas das diferentes relações entre a temperatura e a precipitação que permitiram elaborar índices bioclimáticos. Nestes foi também integrado o tema de evapotranspiração real. O mapa de humidade relativa traduz as disponibilidades hídricas do solo, reflectindo as situações de humidade e de stress hídrico. O mapa é relativo ao período 1961 - 1990 e é disponibilizado em formato raster (resolução espacial de 1x1 km).

### **Modelação de usos de solo e coberto florestal**

A caracterização regional de usos do solo teve por base a sistematização de informação do domínio público, por exemplo o Corine LandCover 1990 e 2000, COS (Carta de Ocupação do Solo) e litologia dos solos com origem no Atlas do ambiente em formato digital. Adicionalmente, utilizou-se informação presente no inventário florestal (cobertura nacional com 129 019 fotopontos) referente ao coberto florestal, nomeadamente montado denso, regular ou disperso de sobro, azinho e misto. Todos os fotopontos presentes em cada freguesia foram tidos em conta para calcular o seu coberto florestal e uso de solo.

## Modelação regional à freguesia

A freguesia foi a base comum a todos os dados, pelo que a informação temática foi “cortada” à freguesia (Ex. Climatologia), enquanto que a informação alfanumérica (Ex. Inventário Florestal) foi agregada à freguesia, juntando-se assim aos dados à freguesia do INE.

## Análise estatística

A metodologia estatística adoptada para a classificação das freguesias pode ser resumida nos três passos seguintes:

- (1) Análise de componentes principais para cada grupo de variáveis, com as 294 freguesias do Alentejo (informação previamente uniformizada no SIG).
- (2) Retenção dos 2 primeiros eixos de cada uma das 6 análises para a classificação das freguesias com K-means, utilizando a informação da ordenação para descrever cada grupo.
- (3) Extrapolação dos questionários através de árvore de classificação (r-part).

Tabela 1 : Variáveis utilizadas para construir os modelos.

Grupo	Abreviatura	Descrição	Máximo	Mínimo
Uso de solo	FL	Floresta	448	0
	AG	Agrícola	293	1
	IC	Inculto	228	0
	SAU	Sup. Agrícola	0,995243	0,121952
	AF	Área Forrageira	0,969761	0
	SA_51	M. misto +denso	2800	0
	SA_50	M. misto denso	1400	0
Tipo de Floresta	SA_30	M. misto disperso	500	0
	SB_51	M. Sobro +denso	11800	0
	SB_50	M. Sobro denso	8500	0
	SB_30	M. Sobro disperso	4000	0
	AZ_51	M. Azinho +denso	8300	0
	AZ_50	M. Azinho denso	10200	0
	AZ_30	M. Azinho disperso	4100	0
	PM_30	PM disperso	1000	0
Litologia	Gra	Granitos	350	0
	Are	Arenitos	497	0
	Meta	Metamórficas	464	0
	Aluv	Aluviões	85	0
	Aluv2	Aluviões 2	39	0
	Casc	Cascalheiras	124	0
	Calc	Calcários	57	0
	Comp	Complexos.....	103	0
Climatologia	TMA	T máx. Agosto	32,5128	26,3425
	Tmj	T mín. Janeiro	8,2308	3,6615
	TMA-mj	Continentalidade	28,6051	18,3816
	Mar	Distância ao mar	21861904	89237,41
	PI	Precipitação Inverno	391,38	189,0534

	PV	Precipitação Verão	55,5508	17,0961	
	PT	Precipitação anual	970,7453	473,7225	
	Alt	Altitude	674,439	22,9898	
	CNF/forr	CN feminino/área forrageira	24,96964	0,0662	
	CNF	CN feminino	1,095486	0,007518	
	CN/forr	CN total/ área forrageira	26,36275	0,103494	
	VL	Vacas leiteiras	3,487612	0	
	OV	Outras vacas	11,45032	0	
	PR	Porcas reprodutoras	3,61768	0	
Gado	OL	Ovelhas leiteiras	7,901152	0	
	OO	Outras ovelhas	25,77678	0,010611	
	CL	Cabras leiteiras	13,43319	0	
	OC	Outras cabras	11,95281	0	
	Bemc	Encabeçamento de Bovinos	11,499	0	
	Semc	Encabeçamento de Suínos	15,74012	0	
	Oemc	Encabeçamento de Ovinos	12,2548	0	
	Cemc	Encabeçamento de Caprinos	5,04246	0	
	Economia: tamanho da herdade	0	Herdade de 0 hectares	0,115385	0
		>0 a <2	Herdade de 0 a 2 hectares	0,837398	0
2 a <4		Herdade de 2 a 4 hectares	0,343434	0	
4 a <8		Herdade de 4 a 8 hectares	0,4	0	
8 a <16		Herdade de 8 a 16 hectares	0,322581	0	
16 a <40		Herdade de 16 a 40 hectares	0,5	0	
40 a <100		Herdade de 40 a 100 Ha	0,428571	0	
>= 100		Herdade com mais de 100H	0,25	0	
Trabalho	UTAr	Unidade de trabalho	3,615862	0,058721	
	UTAF/T	% de trabalho familiar	99,20431	3,814607	
Economia: orientação económica	Cer	ODE cerealicultura	0,819672	0	
	Cad	ODE culturas agrícolas	0,367816	0	
	hort	ODE horticultura	0,465217	0	
	Viti	ODE Vitivinicultura	0,666667	0	
	Frut	ODE Fruticultura	0,32	0	
	Oliv	ODE Olivocultura	0,76	0	
	Cult	ODE Culturas permanentes	0,327434	0	
	BovL	ODE Bovinos de Leite	0,166667	0	
	BovC	ODE Bovinos de Carne	0,410256	0	
	BovLC	ODE Bovinos de Carne e Leite	0,147059	0	
	OvCa	ODE Ovinos e Caprinos	0,625	0	
	Gran	ODE Graníveros	0,22028	0	
	Poli	ODE Policultura	0,247253	0	
	PoIOH	ODE Polipequária Herb	0,210526	0	
	POIOG	ODE Polipequária Gran	0,153846	0	
	AGHer	ODE Agricultura e Herb	0,5	0	
CultDG	ODE Culturas diversas e Gran	0,333333	0		

O objectivo foi numa primeira fase otimizar a escolha das herdades de formar a que estas fossem representativas do ponto de vista multidisciplinar do projecto. No caso Português, utilizou-se o mesmo conjunto de critérios que na Extremadura, mas com a informação uniformizada à freguesia.

Para seleccionar as herdades, optou-se por estudar os diferentes critérios de representatividade ao nível da freguesia. Assim, para cada freguesia do Alentejo calcularam-se (ver Tabela 1):



- (1) Médias climatológicas – Temperatura máxima de Agosto, Temperatura mínima de Janeiro, Continentalidade, Precipitação de Inverno, Precipitação de Verão e Altitude.
- (2) Os usos dominantes – Agrícola, superfície de área útil, área forrageira, Florestal ou Inculto.
- (3) Variáveis económicas - o tamanho médio das herdades, o número de trabalhadores, a orientação económica.
- (4) O tipo de montado presente – Azinho disperso, Azinho denso, Azinho muito denso, Sobre disperso, Sobre denso e Sobre muito denso, Misto disperso, Misto denso e Misto muito denso, Pinheiro manso disperso.
- (5) O tipo de gado presente – Ovinos, Suínos, Bovinos e Caprinos (densidade, nº de herdades e proporção relativa). Efectivo feminino, encabeçamento total.
- (6) Litologia – Arenitos, Granitos, Calcários, Rochas metamórficas e Cascalheiras

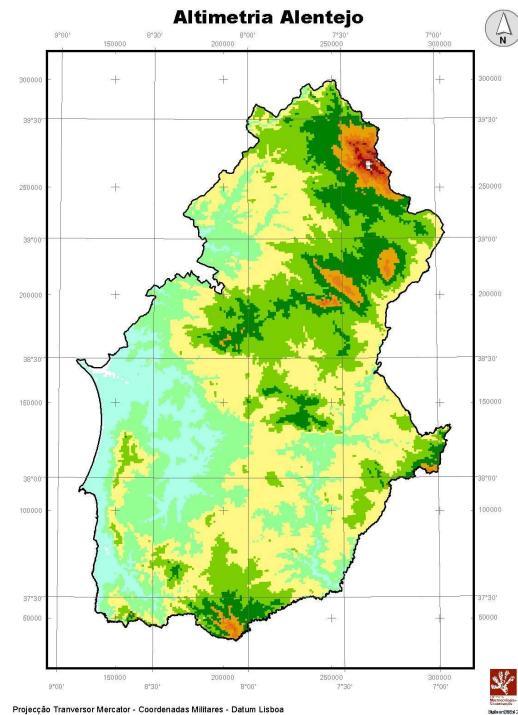
## Resultados

### *Sistema de Informação Geográfica para o Alentejo*

#### Morfologia

O resultado do tratamento e integração da informação geomorfológica no SIG foram os seguintes temas:

- (1) Tema com a altimetria do Alentejo (MDT – Figura 1)
- (2) Tema com a exposição solar
- (3) Tema com os declives
- (4) Rede hidrográfica



**Figura 1** - Modelo Digital Terreno do Alentejo com resolução de 1Km

## Modelação Climática

O resultado do tratamento e integração da informação climatológica no SIG foram os seguintes temas de clima:

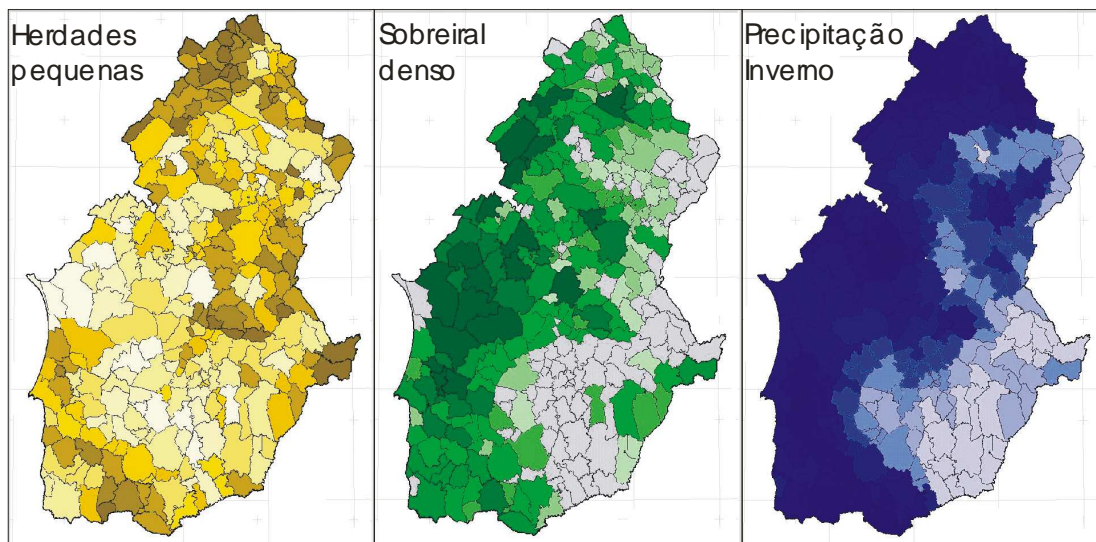
- (4) Temperatura máxima do mês mais quente (Agosto).
- (5) Temperatura mínima do mês mais frio (Janeiro).
- (6) Precipitação média mensal dos meses de Janeiro e Agosto.
- (7) Precipitação anual.

Os temas climáticos preparados referem-se ao conjunto de variáveis de síntese para a temperatura e precipitação (média mensal e anual).

## Modelação bioclimática

O resultado do tratamento e integração da informação climatológica no SIG foram os seguintes temas bioclimáticos (não usados nos modelos estatísticos):

- (8) Tema com o Índice de Aridez
- (9) Vários temas de Índices bioclimáticos
- (10) Tema da evapotranspiração real
- (11) Tema de humidade relativa



**Figura 2** - Informação uniformizada à freguesia, para dados alfa-numéricos (Herdades pequenas – dados do INE à freguesia; Sobreiral denso – dados sumarizados a partir do inventário florestal) e informação temática (Precipitação de Inverno – tema de informação contínua recortado à freguesia). Para cada variável, os tons mais escuros correspondem a valores mais altos.

## Modelação Usos do solo

O resultado do tratamento e integração da informação climatológica no SIG foram os seguintes temas bioclimáticos (não usados nos modelos estatísticos):

- (12) Tema com a litologia regional
- (13) Tema com os dois períodos de referência no Corine LandCover
- (14) Temas regionais da carta de ocupação do solo
- (15) Temas regionais de coberto florestal

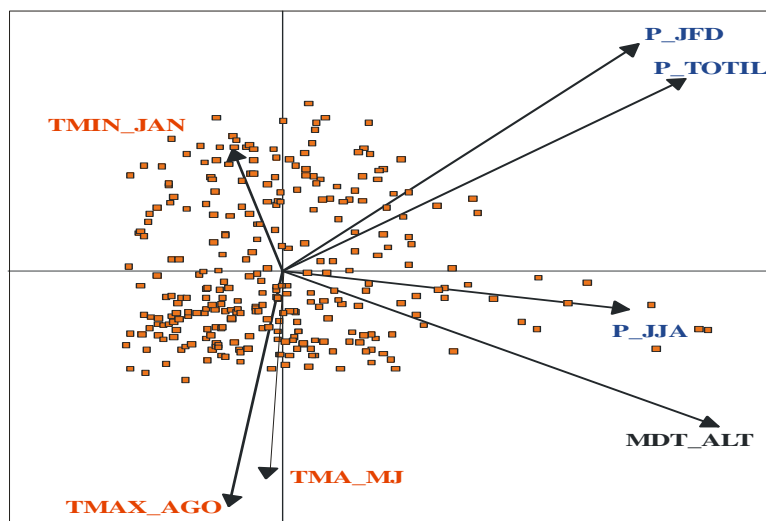
### **Regionalização e uniformização dos dados**

O resultado da regionalização dos dados está ilustrado na figura 2; a Freguesia passou a ser o mínimo denominador comum de toda a informação compilada no SIG, seja ela contínua (Ex. Precipitação de Inverno), alfanumérica (Ex. Sobreiral denso) ou presente numa tabela do INE (Ex. quantidade de herdades pequenas). Esta informação pode extraída e facilmente trabalhada fora do SIG. Os resultados do tratamento estatístico (ordenações, classificações e extrapolações) podem ser facilmente visualizados através do SIG regional.

### **Ordenação, classificação do Alentejo e extrapolação de dados**

#### **Ordenação: análise de componentes principais**

Em resultado das Análises de componentes principais (ACP) resultaram 12 variáveis, correspondentes aos 2 primeiros eixos da biplot para cada grupo de variáveis. Apesar de se perder informação, a informação essencial ficou retida nestes eixos. Por Ex., no caso da Climatologia (Figura 3), as variáveis relacionadas com a precipitação ficaram retidas no primeiro eixo enquanto que as variáveis de temperatura ficaram projectadas no segundo eixo.



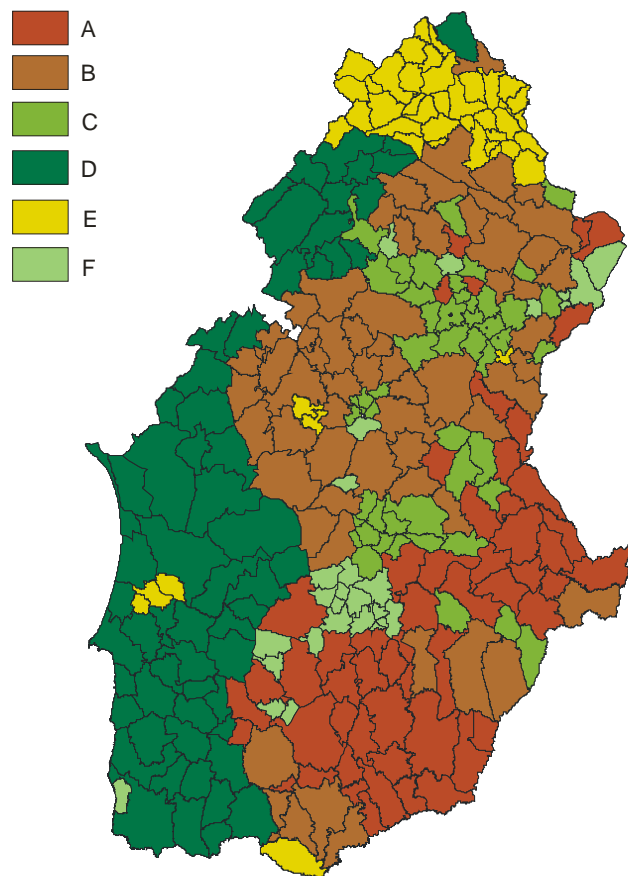
**Figura 3** - ACP com as variáveis climatológicas. A precipitação (variáveis a azul) ficou representada no primeiro eixo enquanto que a temperatura (variáveis a laranja) ficaram representadas pelo 2º eixo.

Para cada grupo de variáveis obtivemos uma ACP que foi posteriormente utilizada na classificação.

#### **Classificação das 294 Freguesias do Alentejo**

Através do algoritmo *K-means*, as freguesias foram classificadas de acordo com descritores simplificados (os dois primeiros eixos) dos grupos de variáveis em 6 classes. A classificação obtida foi analisada de forma a determinar que variáveis caracterizavam cada classe (ver discussão).

As 6 classes de freguesias obtidas são bastante homogéneas, sendo as 12 variáveis utilizadas (1º e 2º eixos das PCA de cada um dos 6 grupos de variáveis) muito significativas para a classificação final ( $p < 0,001$ ).



**Figura 4** – Classificação final obtida. Os 6 grupos (A a F) estão descritos no texto

### **Estratificação da escolha das herdades**

Com base nesta classificação, foram escolhidas as 31 herdades a estudar no Alentejo.

Para escolher estas herdades, foram seguidas três regras:

1º Para cada tipo de montado teriam que ser escolhidas pelo menos três herdades.

2º O número de herdades será proporcional à área de montado de cada tipo.

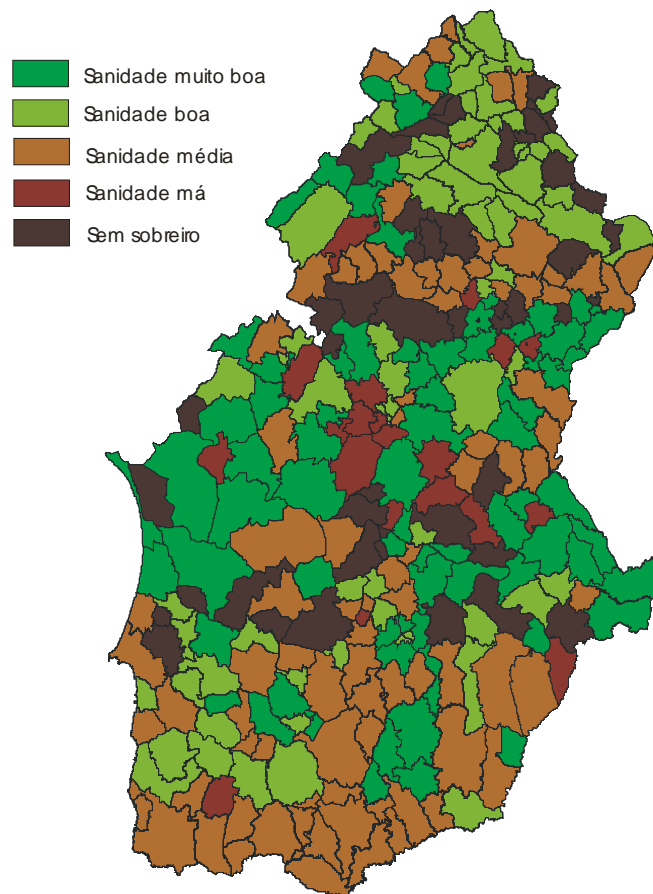
3º Para o tamanho da herdade (importante variável económica) é necessário garantir na escolha das herdades a distribuição de tamanhos desejada pelo grupo da Economia, já que à freguesia esta variável é muito heterogénea.

Obedecendo a estas três regras, as herdades foram escolhidas para todo o Alentejo, obtendo respectivamente:

- (1) Grupo A – 5 herdades,
- (2) Grupo B – 7 herdades
- (3) Grupo C – 4 herdades
- (4) Grupo D – 7 herdades
- (5) Grupo E – 3 herdades
- (6) Grupo F – 3 herdades

### Extrapolação dos dados recolhidos nos inquéritos

Através dos inquéritos obtivemos mais de 300 novas variáveis que podemos agora extrapolar para todo o Alentejo. Mais de 20 variáveis foram extrapoladas, relacionadas com o uso agrícola, factores de sanidade, SAU, UTA, tipo de gado e factores económicos. Como exemplo apresentamos o resultado da extrapolação da sanidade do Sobreiro. As áreas de boa e má sanidade resultam de duas combinações independentes de variáveis, sendo por isso independentes da quantidade de Sobreiro existente.



**Figura 5** - Resultado da extrapolação dos questionários relativa à sanidade do montado de sobreiro. A verde estão áreas sem problemas, a castanho, áreas com problemas de sanidade, castanho escuro são freguesias que praticamente não têm montado de sobreiro.

Concretizando, em Barrancos não há muito Sobreiro, mas o que existe não sofre de problemas sanitários. Analogamente, a freguesia de Alcácer do Sal também tem uma boa sanidade, mas é uma zona de algum Sobreiro. Em zonas de elevada concentração de sobreiro, podem existir situações de boa sanidade (Ex. Montargil) ou de má sanidade (Ex.

Aldeia Velha, no conselho de Avis). Em zonas de fraca concentração de sobro, a região de Évora é uma zona de baixa sanidade enquanto que a zona de Barrancos apresenta bons índices de sanidade. Este tipo de resultados pode ser obtido para qualquer variável recolhida ao escala da herdade, desde que uma parte importante da sua variância seja explicada pelas variáveis da Tabela 1.

## Discussão

A análise de componentes principais (*ACP*) permitiu reduzir a matriz de setenta e duas variáveis (ver Tabela 1) numa matriz de doze variáveis correspondente aos dois primeiros eixos de cada *ACP* (uma *ACP* por grupo de variáveis). Adicionalmente, a redução da informação grupo a grupo permitiu dar o mesmo peso a cada grupo de variáveis na classificação final obtida. A informação das 6 *ACP* pode ser cruzada com a dos grupos resultantes da classificação das freguesias, de forma a tipificar os grupos obtidos.

(1) A climatologia produz uma *ACP* bastante homogénea, com a Temperatura mínima de Janeiro a discriminar um grupo de freguesias oceânicas (agrupadas no grupo D). Já as precipitações elevadas ocorrem essencialmente nas freguesias do grupo E (Norte Alentejano). As freguesias de maior continentalidade agrupam-se nos grupos A, C e F, enquanto que o grupo B é intermédio entre o continental e o oceânico. (2) Para os Usos de solo, a *ACP* discrimina bem as Freguesias essencialmente florestais (agrupadas no grupo B e D) das restantes. As freguesias essencialmente agrícolas, aparecem agrupadas no grupo A. (3) Quanto à litologia, o Alentejo é claramente dividido em três partes distintas: uma dominada por arenitos (grupo D, no litoral), uma segunda dominada por granitos (grupo B, E e F, mais no interior) e outra dominada por metamórficas (grupo A e C). (4) Para a cobertura florestal, a *ACP* com as variáveis da floresta separam claramente as freguesias dominadas por montado de Sobro, das freguesias dominadas por montado de Azinho; a densidade do montado, á freguesia, é uma variável não discriminante. (5) O gado é bastante variado nas freguesias, pelo que é um descritor fraco dos diferentes tipos de montado. Apenas o grupo de freguesias do grupo C parecem estar relacionadas com uma presença elevada de ovelhas, suínos e bovinos enquanto que o grupo E está relacionado com a presença de cabras; o grupo B tem vacas e o grupo D e A não têm praticamente gado. (6) As freguesias dos grupos D, A e F são caracterizadas por terem herdades pequenas ou médias. Em oposição, as herdades grandes estão concentradas nas freguesias do grupo B. O grupo E tem as herdades mais pequenas (courelas) e o maior número de trabalhadores por hectare.

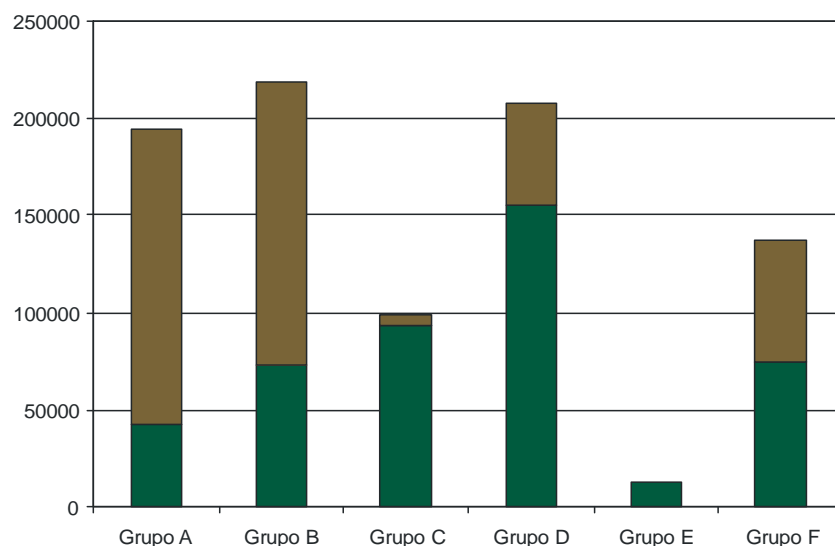
Deste cruzamento de informação podemos agora tipificar a classificação da seguinte forma:

- (1) **Grupo A** – Agrícola, Azinhal, sem Sobreiral, Granito, Precipitação muito baixa, UTA muito baixo, Herdades pequenas.
- (2) **Grupo B** – Sau elevado, incultos, Azinhal, algum Sobreiral, Granito, Temperatura alta, Precipitação média, Muitas vacas, Herdades grandes, Bovinos de carne, Ovinos e Caprinos.
- (3) **Grupo C** – Agrícola, Montado misto, Litologia variada, Temperatura alta, Muitos Ovinos e Bovinos, UTA alto, Olival.
- (4) **Grupo D** – Florestal, Sobreiral, Arenitos, Precipitação boa, pouco gado (alguns bovinos carne), Herdades grandes, Hortas.

- (5) **Grupo E** –Boa Forrageira, Pouco Sobreiral, Metamórficas, Sem Arenitos, Boa precipitação, Vacas leiteiras, Herdades pequenas, UTA familiar alto, Bovinos de carne.
- (6) **Grupo F** – Agrícola não florestal, Pouco Sobreiral, Metamórficas sem arenitos, Temperatura alta, Baixa precipitação, Sem suínos, UTA alto, Cerealífero.

A contiguidade espacial dos agrupamentos de freguesias (ver figura 4) sugerem uma classificação coerente e concordante com outros trabalhos anteriores (Pereira e Fonseca 2003). A concordância espacial dos factores antropogénicos e naturais pode ser explicada pela antiguidade da origem do montado, como acontece com várias espécies de árvores na Europa (Svenning & Skov 2005).

A validação da escolha das herdades foi feita à posteriori, comparando os atributos de cada um dos 6 grupos obtidos através da classificação. Na Figura 6 está representado o coberto efectivo de sobre e azinho para todas as herdades escolhidas em cada grupo. Comparando estes resultados com a tipificação dos grupos, vemos que a escolha das herdades foi eficaz, já que o grupo D revelou ser aquele com mais montado de Sobre, o grupo A e B com mais montado de Azinho e o grupo C, E e F caracterizados por serem menos florestais (mas com a quantidade relativa de sobre e azinho também concordantes com a classificação obtida)



**Figura 6** - A verde está representada a cobertura média de sobre nas herdades pertencentes a cada grupo classificado. A castanho está representada a cobertura de azinho.

A extrapolação dos dados recolhidas para cada herdade para todo o Alentejo é possível através da utilização da matriz de dados inicial (Tabela 1); no entanto, apenas para as variáveis que temos informação detalhada a extrapolação tem um resultado interessante. As variáveis que não estão relacionadas com as variáveis da Tabela 1 não poderão ser modelizadas, já que nunca poderá ser obtida uma extrapolação significativa.

Estes primeiros resultados são extremamente promissores, já que a grande vantagem deste método é poder obter informação muito detalhada à escala do Alentejo que por outra via seria demasiado cara de obter ou demasiado morosa. A metodologia pode ser detalhada e a rede de herdades pode ser alargada de forma a ter graus de liberdade suficientes para uma extrapolação mais fiável, já que o principal problema encontrado tem a ver com o reduzido número de herdades, que poderá desta forma ser minimizado.

O impacto dos factores antropogénicos nos ecossistemas há muito que é um dado adquirido num mundo crescentemente humanizado (McLaughlin & Mineau 1995, Tappeiner et al. 1998, Pereira & da Fonseca 2003), nomeadamente o impacto do pastoreio (Reed & Clokie 2000, Watkinson & Ormerod 2001) e dos factores associados ao tipo de exploração do montado (Pereira & da Fonseca 2003) .

Por isso, a aproximação multidisciplinar do Montado, integrando factores humanos e naturais é sem dúvida a forma adequada de tratar este complexo ecossistema que resulta da interacção entre o Homem e a Natureza (Pereira & Fonseca 2003); apenas a manutenção da biodiversidade permitirá a manutenção da produtividade dos ecossistemas extensivos (McLaughlin & Mineau 1995, Tilman 2000), enquanto que o seu decréscimo terá consequências dramáticas no funcionamento, estabilidade e sustentabilidade deste interessante ecossistema (Tilman et al. 1996).

## Bibliografia

- Balmford, A., & W. Bond. 2005. Trends in the state of nature and their implications for human well-being. *Ecology Letters* 8:1218-1234.
- Blondel, J. & Aronson, J. 1999. *Biology and Wildlife of the Mediterranean Region*. Oxford University Press Inc., New York.
- Breiman, L., Friedman, J. H., Olshen, R. A. & Stone, C. J. 1984. *Classification and Regression Trees*. London: Chapman and Hall.
- Carrión, J. S., I. Parra, C. Navarro, & M. Munera. 2000. Past distribution and ecology of the cork oak (*Quercus suber*) in the Iberian Peninsula: a pollen-analytical approach. *Diversity and Distributions* 6:29-44.
- Clark, L. A. & Pregibon, D. 1990. Tree-based models. In *Statistical Models in S*, eds. J. M. Chambers and T. J. Hastie. Pacific Grove, Calif.: Wadsworth & Brooks/Cole
- Coelho, I. S. 1996. O Montado, a Economia e o Desenvolvimento do Alentejo. *Silva Lusitana* 4:39-48
- De Miguel, J. M. 1999. Naturaleza y configuración del paisaje agrosilvopastoral en la conservación de la diversidad biológica en España. *Revista Chilena de Historia Natural* 72:547-557.
- Gahegan, M., Wachowicz, M., Harrower, M., & Rhyne, T. M. 2000 The Integration of Geographic Visualization with Knowledge Discovery in Databases and Geocomputation. *ICA Commission on Visualization: Working Group on Database-Visualization Links*
- Goodchild M.F., Parks, B. O. & Steyaert, L. T. 1993. *Environmental modeling with GIS*. Oxford University Press, New York, Oxford.
- Haslett, J. R. 1990. Geographic information systems: a new approach to habitat definition and the study of distributions. *TREE*. 5, 214-218.
- Herzog, F. 1998. Streubost: a traditional agroforestry system as a model for agroforestry development in temperate Europe. *Agroforestry Systems* 42:61-80.
- Joffre, R., S. Rambal, & J. P. Ratte. 1999. The dehesa system of southern Spain and Portugal as a natural ecosystem mimic. *Agroforestry Systems* 45:57-79.
- Legendre, P. & Legendre, L. 1998. *Numerical Ecology (Second English Edition)*. Elsevier, Amsterdam.
- Leiva M. J., F. S. Chapin, & R. F. Ales. 1997. Differences in species composition and diversity among Mediterranean grasslands with different history—the case of California and Spain. *Ecography* 20:97-106.



- McLaughlin, A., & P. Mineau. 1995. The impact of agricultural practices on biodiversity. *Agriculture Ecosystems & Environment* **55**:201-212.
- Nyerges, T. L. 1993. Understanding the scope of GIS: its relationship to environmental modeling. In *Environmental Modeling with GIS*. eds. Goodchild, M. F., Parks, B. O. & Steyaert, L. T. pp. 75-93. Oxford University Press, New York.
- Onipchenko, V. G., & G. V. Semenova. 1995. Comparative-analysis of the floristic richness of alpine communities in the caucasus and the central alps. *Journal of Vegetation Science* **6**: 299–304.
- Pereira, P.M. & Pires da Fonseca, M. 2003 Nature vs. nurture: the making of the montado ecosystem. *Conservation Ecology*, 7, 7 [Online].
- Pinto-Correia, T. 1993. Threatened landscape in Alentejo, Portugal: the montado and other agro-sylvo-pastoral systems. *Landscape and Urban Planning* **24**:43–48.
- Prentice, I. C. 1977. Non-metric ordination methods in ecology. *Journal of Ecology*. **65**, 85-94.
- Reed, M.S. & Clokie, M.R.J. 2000 Effects of grazing and cultivation on forest plant communities in Mount Elgon National Park, Uganda. *African Journal of Ecology*, **38**, 154-162.
- Santini, F. & Angulo, A. 2001. Assessing conservation biology priorities through the development of biodiversity indicators. *Rivista di Biologia/Biology Forum*. **94**, 259-276.
- Svenning, J.-C., & F. Skov. 2005. The relative roles of environment and history as controls of tree species composition and richness in Europe. *Journal of Biogeography* **32**:1019-1033.
- Tappeiner, U., Tasser, E., & Tappeiner, G. 1998 Modelling vegetation patterns using natural and anthropogenic influence factors: Preliminary experience with a GIS based model applied to an Alpine area. *Ecological Modelling*, **113**, 225-237.
- Ter Braak, C. J. F. & Looman, C. W. N. 1994. Biplots in reduced-rank regression. *Biometrical Journal*. **36**, 983-1003.
- Tilman, D. 2000 Causes, consequences and ethics of biodiversity. *Nature*, **405**, 208-211.
- Tilman, D., Wedin, D., & Knops, J. 1996 Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature*, **379**, 718-720.
- Walker, P. A. 1990. Modeling wildlife distributions using a geographic information system: kangaroos in relation to climate. *Journal of Biogeography*. **17**, 279-289.
- Watkinson, A.R. & Ormerod, S.J. 2001 Grasslands, grazing and biodiversity: editors' introduction. *Journal of Applied Ecology*, **38**, 233-237.
- Woodhouse, S. P., J. E. G. Good, A. A. Lovett, R. J. Fuller, & P. M. Dolman. 2005. Effects of land-use and agricultural management on birds of marginal farmland: a case study in the Llyn peninsula, Wales. *Agriculture Ecosystems & Environment* **107**:331-340.