



CENTRO INTERNACIONAL DE ALTOS ESTUDIOS AGRONOMICOS MEDITERRANEOS

INSTITUTO AGRONOMICO MEDITERRANEO DE ZARAGOZA

**ANÁLISE DA ESTRUTURA ECOLÓGICA
DE TRÁS-OS-MONTES E ALTO DOURO
UTILIZAÇÃO DE MÉTODOS DE ANÁLISE MULTIVARIADA
E TEORIA DA INFORMAÇÃO**

José Manuel CORREIA DOS SANTOS FERREIRA DE CASTRO

**TESIS PRESENTADA Y PUBLICAMENTE
DEFENDIDA EN EL I.A.M.Z. PARA
LA OBTENCION DEL DIPLOMA DE
ALTOS ESTUDIOS DEL C.I.H.E.A.M.**

MASTER OF SCIENCE

Zaragoza, Marzo 1996

**ANALISE DA ESTRUTURA ECOLOGICA
DE TRAS-OS-MONTES E ALTO DOURO
UTILIZAÇÃO DE METODOS DE ANALISE MULTIVARIADA
E TEORIA DA INFORMAÇÃO**

José Manuel CORREIA DOS SANTOS FERREIRA DE CASTRO

Zaragoza, Marzo 1996

CENTRO INTERNACIONAL DE ALTOS ESTUDIOS AGRONOMICOS MEDITERRANEOS

INSTITUTO AGRONOMICO MEDITERRANEO DE ZARAGOZA

ANALISE DA ESTRUTURA ECOLOGICA DE TRAS-OS-MONTES E ALTO DOURO

UTILIZAÇÃO DE METODOS DE ANALISE MULTIVARIADA

E TEORIA DA INFORMAÇÃO

José Manuel CORREIA DOS SANTOS FERREIRA DE CASTRO

Trabajo realizado bajo la dirección conjunta del Dr. Antonio GOMEZ SAL del Instituto Pirenaico de Ecología, Consejo Superior de Investigaciones Científicas, Zaragoza, y del Dr. João Manuel SANTOS BENTOS del Departamento Forestal, Universidad de Tras-Os-Montes e Alto Douro, Vila Real,

y presentado en lectura pública el día 7 de Marzo de 1996, ante el siguiente tribunal :

- Antonio ESCARRRE ESTEVE, Departamento de Ecología, Facultad de Ciencias, Universidad de Alicante,
- Miguel MOREY ANDREU, Area de Ecología, Facultad de Ciencias, Universidad de las Islas Baleares, Palma de Mallorca,
- Francisco CASTRO REGO, Departamento de Estadística, Instituto Superior de Agronomía, Universidad Técnica de Lisboa,
- Francisco DIAZ PINEDA, Departamento de Ecología, Facultad de Biología, Universidad Complutense de Madrid,
- Dunixi GABIÑA ITURRIAGA, Director Adjunto del Instituto Agronómico Mediterráneo de Zaragoza.

À Marina e ao Eduardo

AGRADECIMENTOS

Ao Excelentíssimo Presidente do Instituto Politécnico de Bragança, Professor Doutor DIONÍSIO AFONSO GONÇALVES, aos Excelentíssimos Directores da Escola Superior Agrária de Bragança, Professor Doutor FRANCISCO TERROSO CEPEDA e Professor Doutor ALFREDO JORGE DA COSTA TEIXEIRA, expressamos o nosso reconhecimento pela disponibilização dos meios necessários à nossa promoção académica e valorização profissional.

Ao Excelentíssimo Director do Instituto Agronómico Mediterrâneo de Saragoça, D. MIGUEL VALLS agradecemos a oportunidade dada para a realização do curso de pós-graduação e desta prova.

Ao Professor Doutor ANTONIO GÓMEZ SAL e Professor Doutor JOÃO SANTOS BENTO, expressamos a nossa gratidão pela sábia orientação científica e excepcional disponibilidade demonstrada para com a realização do presente trabalho.

Ao Professor Doutor FRANCISCO CASTRO REGO, agradecemos a disponibilidade para a discussão dos métodos estatísticos abordados.

Ao Professor Doutor RUI MANUEL CORTES e ao Engenheiro CARLOS GONÇALVES AGUIAR, pela disponibilidade dos programas informáticos para a realização da análise matemática dos dados do trabalho.

Ao Engenheiro TOMÁS DE AQUINO FIGUEIREDO, agradecemos o apoio na consulta e utilização dos dados relativos à Carta dos Solos, de Uso do Solo e Aptidão do Solo do Nordeste.

À Dra. MARIA JOSÉ MIRANDA e ao Dr. JOÃO SÉRGIO SOUSA, agradecemos o apoio às traduções para espanhol, francês e inglês.

Ao ATILANO SUAREZ agradecemos a disponibilidade e empenho posto na realização da composição gráfica e impressão do texto e seus anexos.

Ao Dr. JOÃO SÉRGIO SOUSA e Dr. PEDRO SALGUEIRO, agradecemos o seu total apoio e permanente incentivo para a conclusão deste trabalho.

A todos que, não sendo referidos antes, de uma forma directa ou indirecta permitiram a realização e conclusão deste trabalho, o nosso muito obrigado.

ÍNDICE GERAL

AGRADECIMENTOS	iii
ÍNDICE GERAL	iv
ÍNDICE DE FIGURAS.....	vi
ÍNDICE DE QUADROS.....	viii
ÍNDICE DE ANEXOS	x
RESUMO	xi
RESUMÉN	xii
RÉSUMÉ.....	xiii
ABSTRACT	xiv
CAPÍTULO 1 INTRODUÇÃO	1
CAPÍTULO 2 REVISÃO DE METODOLOGIAS PARA ANÁLISE DE DADOS.....	7
2.1 Análise de Complexos de Dados Ecológicos.....	7
2.2 Breve Resenha de Métodos de Análise Multivariada	9
2.2.1 MÉTODOS DE CLASSIFICAÇÃO.....	11
2.2.1.1 MÉTODOS HIERÁRQUICOS AGLOMERATIVOS.....	12
<i>optimização da distância</i> 12	
<i>optimização da homogeneidade</i> 13	
<i>critério da teoria da informação</i> 13	
<i>optimização global</i> 14	
2.2.1.2 MÉTODOS HIERÁRQUICOS DIVISIVOS.....	15
<i>associação</i> 15	
<i>"twinspan"</i> 16	
2.2.1.3 MÉTODOS NÃO HIERÁRQUICOS	17
<i>optimização global</i> 17	
<i>"c-médias" indistintas</i> 18	
<i>árvore de menor distância</i> 18	
2.2.2 MÉTODOS DE ORDENAÇÃO	18
2.2.2.1 ANÁLISES INDIRECTAS DE GRADIENTES	20
<i>análise de componentes principais</i> 20	
<i>análise factorial</i> 22	
<i>análise de correspondências</i> 23	
<i>análise de coordenadas principais</i> 24	
2.2.2.2 ANÁLISE DIRECTA DE GRADIENTES.....	25
<i>análise das correlações canónicas</i> 25	

2.3 Entropia e Informação: significado e utilização.....	29
CAPÍTULO 3 ECOLOGIA DE TRÁS-OS-MONTES E ALTO DOURO	35
3.1 Apresentação do território.....	35
3.2 Apresentação dos factores estudados	38
3.2.1 VARIÁVEIS CLIMÁTICAS	38
3.2.2 VARIÁVEIS FISIGRÁFICAS	42
3.2.3 VARIÁVEIS DE USO E OCUPAÇÃO DO SOLO.....	43
3.3 Metodologia utilizada	44
3.3.1 CLASSIFICAÇÃO HIERÁRQUICA DIVISIVA.....	44
3.3.2 ANÁLISE DE CONCENTRAÇÃO.....	54
3.3.3 ANÁLISE PELA TEORIA DA INFORMAÇÃO	55
3.4 Comparação dos resultados com outras classificações	58
3.4.1 GRANDES REGIÕES (AGROCONSULTORES E COBA, 1990)	59
3.4.2 CARTA ECOLÓGICA DE PORTUGAL (ALBUQUERQUE, 1945).....	62
3.4.3 CARTA DE AGROTIPOS DE PORTUGAL (ALBUQUERQUE, 1945)	67
CAPÍTULO 4 CONCLUSÕES	71
4.1 Considerações Finais.....	71
4.2 Principais rasgos da ecologia de Trás-os-Montes e Alto Douro	74
SÍNTESE	79
SINTESIS	87
SYNTHÈSE.....	95
SYNTHESIS.....	103
BIBLIOGRAFIA.....	110
ANEXOS.....	A

ÍNDICE DE FIGURAS

Figura 2.1- Síntese das contribuições de outras ciências à investigação em ciências do meio natural (adapt. Legendre e Legendre, 1984a)	7
Figura 2.2 - Estruturação conceptual de dados ecológicos (adapt. Legendre e Legendre, 1984a).....	8
Figura 2.3 - Representação esquemática de um espaço (I), com o respectivo modelo de classificação (II) e ordenação (III), tendo em conta 9 atributos por objecto, com 4 classes cada () (adapt. Zonneveld, <i>in press</i>).	9
Figura 2.4 - Representação das distâncias de decisão para atribuição do objecto "o" a três grupos diferentes (vizinho mais próximo, centróide e elemento mais afastado).....	13
Figura 2.5 - Desenvolvimento-base da análise de componentes principais (modelo com utilização da matriz de correlação).....	21
Figura 2.6 - Desenvolvimento-base da análise factorial.....	22
Figura 2.7 - Desenvolvimento-base da análise de correspondências.....	23
Figura 2.8 - Desenvolvimento-base da análise de coordenadas principais.....	24
Figura 2.9 - Desenvolvimento-base da análise das correlações canónicas.....	25
Figura 2.10 - Desenvolvimento-base da análise discriminante.....	27
Figura 2.11 - Decomposição da informação em decisões binárias entre 4 símbolos (N) (adapt. Naveh e Lieberman, 1983)	30
Figura 2.12 - Representação gráfica da relação entre os conceitos de entropia e informação mútua (adapt. Daget e Godron, 1982)	31
Figura 2.13 - Informação interactiva (área sombreada à esquerda) e informação mútua (área sombreada à direita) de uma distribuição de frequência com três dimensões. (Orlóci, 1991a)	32
Figura 3.1 - Dendrograma relativo às partições realizadas pela metodologia do "Twinspan"	49
Figura 3.2 - Representação gráfica dos grupos de locais e grupos de atributos no espaço definido pelos dois primeiros eixos canónicos	55

Figura 3.3 - Representação gráfica dos grupos de locais definidos pelo "Twinspan" e as classes da classificação das Grandes Regiões (Agroconsultores e COBA, 1990), no espaço definido pelos dois primeiros eixos canónicos	61
Figura 3.4 - Representação gráfica dos grupos de locais definidos pelo "Twinspan" e as classes da classificação da Carta Ecológica (Albuquerque, 1945), no espaço definido pelos dois primeiros eixos canónicos.....	66
Figura 3.5 - Representação gráfica dos grupos de locais definidos pelo "Twinspan" e as classes da classificação da Carta de Agrotipos (Albuquerque, 1945), no espaço definido pelos dois primeiros eixos canónicos.....	69
Figura 4.1- Representação gráfica dos valores do índice de Sørensen de dissimilaridade entre a classificação do "Twinspan" e outras com as quais é comparada, para cada meridiano considerado na recolha de dados.....	77

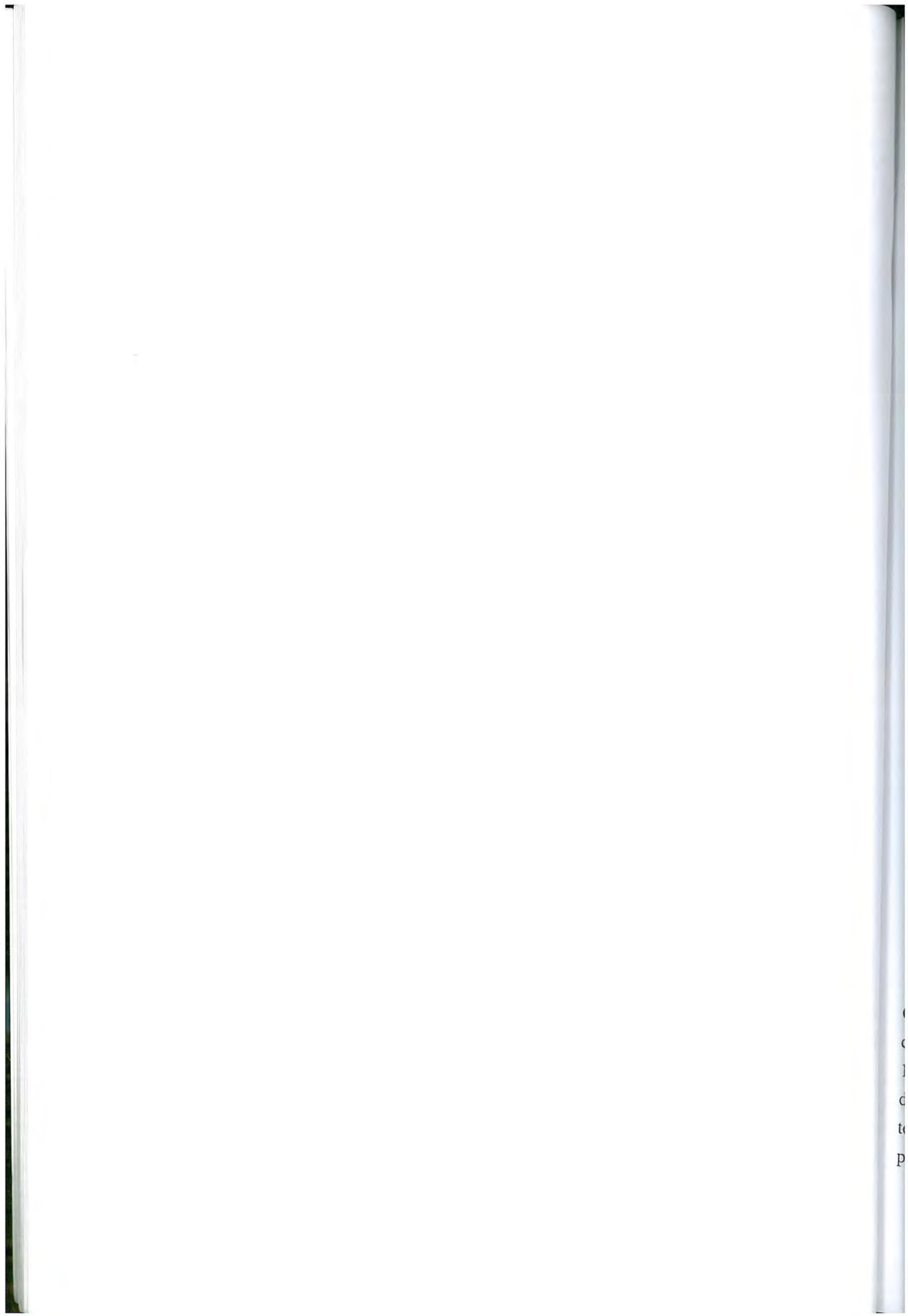
ÍNDICE DE QUADROS

Quadro III.1 - Relação dos pontos retirados da análise, respectivo ponto mais semelhante e valor da sua dissemelhança.....	37
Quadro III.2 - Variáveis quantitativas utilizadas no estudo, com representação cartográfica do tipo isolinha.	39
Quadro III.3 - Variáveis quantitativas e qualitativas utilizadas no estudo com representação cartográfica do tipo área.	41
Quadro III.4 - Variáveis preferenciais correspondentes à divisão entre os grupos 0 e 1.....	49
Quadro III.5 - Variáveis preferenciais correspondentes à divisão entre os grupos 00 e 01.	50
Quadro III.6 - Variáveis preferenciais correspondentes à divisão entre os grupos 10 e 11.	51
Quadro III.7 - Variáveis preferenciais correspondentes à divisão entre os grupos 000 e 001.	51
Quadro III.8 - Variáveis preferenciais correspondentes à divisão entre os grupos 010 e 011.	52
Quadro III.9 - Variáveis preferenciais correspondentes à divisão entre os grupos 100 e 101.	52
Quadro III.10 - Variáveis preferenciais correspondentes à divisão entre os grupos 110 e 111.	52
Quadro III.11 - Classificação dos atributos utilizados no estudo.	53
Quadro III.12 - Tabela de contingência entre grupos de locais e grupos de variáveis definidos.	54
Quadro III.13 - Teste de χ^2 e coeficientes de correlação canónica entre grupos de locais e grupos de atributos.....	54
Quadro III.14 - Ordenação dos atributos pelo seu valor em informação mútua.....	56
Quadro III.15 - Valores de "C" mais elevados para os atributos climáticos, fisiográficos e de uso do solo, na classificação realizada com "Twinspan ".	58
Quadro III.16 - Tabela de contingência entre os grupos de locais definidos pelo "Twinspan" e as classes da classificação das Grandes Regiões.	60

Quadro III.17 - Teste de χ^2 e coeficientes de correlação canónica entre grupos de locais definidos pelo "Twinspan" e a classificação das Grandes Regiões.	61
Quadro III.18 - Valores de "C" mais elevados para os atributos climáticos, fisiográficos e de uso do solo, na sectorização realizada pela classificação das Grandes Regiões.	62
Quadro III.19 - Tabela de contingência entre os grupos de locais definidos pelo "Twinspan" e as classes da classificação da Carta Ecológica.	65
Quadro III.20 - Teste de χ^2 e coeficientes de correlação canónica entre grupos de locais definidos pelo "Twinspan" e a classificação da Carta Ecológica	65
Quadro III.21 - Valores de "C" mais elevados para os atributos climáticos, fisiográficos e de uso do solo, na sectorização realizada pela Carta Ecológica.	66
Quadro III.22 - Tabela de contingência entre os grupos de locais definidos pelo "Twinspan" e as classes da classificação da Carta de Agrotipos.	68
Quadro III.23 - Teste de χ^2 e coeficientes de correlação canónica entre grupos de locais definidos pelo "Twinspan" e a classificação da Carta de Agrotipos.	68
Quadro III.24 - Valores de "C" mais elevados para os atributos climáticos, fisiográficos e de uso do solo, na sectorização realizada pela Carta de Agrotipos.	70
Quadro IV.1 - Coeficientes de dissimilaridade entre as 4 classificações em confronto (coeficientes de Sørensen e de Ochiai).	76
Quadro IV.2 - Coeficientes de dissimilaridade inferiores a 0.5 entre as classes da classificação do "Twinspan" e as classes das outras classificações em confronto.	78

ÍNDICE DE ANEXOS

Anexo 1	B
- Localização aproximada dos meridianos e locais de amostragem.....	C
Anexo 2	D
- Tabela de dupla entrada resultante da análise de "Twinspan"	E
Anexo 3	F
- Mapa da Temperatura Média Anual	G
- Mapa da Precipitação Total Anual.....	H
- Mapa Geomorfológico	I
- Mapa Hipsométrico	J
- Mapa de Distribuição das Culturas Arvenses de Sequeiro	K
- Mapa de Distribuição de Carvalhos e Castanheiro	L
Anexo 4	M
- Classificação das Grandes Regiões.....	N
- Classificação da Carta Ecológica.....	O
- Classificação da Carta de Agrotipos.....	P



RESUMO

O presente trabalho baseia-se na utilização de métodos de análise multivariada e da teoria da informação no estudo da ecologia de Trás-os-Montes e Alto Douro. Na primeira parte discutem-se as principais metodologias multivariadas desenvolvidas até hoje, na segunda testam-se as possibilidades de aplicação de uma técnica simples de amostragem sistemática das principais variáveis ambientais, numa escala territorialmente ampla. Os resultados foram confrontados com anteriores sectorizações de carácter sintético, de âmbito regional ou nacional, descritas para Trás-os-Montes e Alto Douro. A informação ambiental, do tipo climático, fisiográfico e de uso do solo, foi recolhida de diversa bibliografia publicada e cartografada sobre o tema, para pontos de amostragem em cada 1 minuto de latitude, de meridianos espaçados 15 minutos de longitude. Realizaram-se análises de classificação, de concentração e de informação. A técnica "TWINSPAN" de classificação separou claramente os ambientes de influência atlântica dos de influência continental, assumindo grande importância nessa discriminação os factores climáticos característicos dos climas oceânicos. Os factores do tipo fisiográfico e de uso de solo assumem maior relevo na diferenciação mais detalhada dos oito sectores estabelecidos neste estudo. Nos sectores atlânticos, individualizam-se os ambientes das cumeadas graníticas de utilização florestal e pastoril, os ambientes das comunidades rurais com a sua paisagem agrária de autoconsumo, a montanha média de origem xistosa e utilização agrícola de sequeiro, e as zonas de ribeira, com largos socacos e culturas de regadio. Nos sectores de menor influência atlântica, destacam-se a paisagem vitícola do Douro, os vales dos seus afluentes, de utilização fundamentalmente cerealífera e frutícola, o planalto predominantemente cerealífero com "lameiros de secadal", e as elevações graníticas, sem qualquer utilização agrária dominante aparte de pequenos espaços com vinha e pinhal. A análise de concentração detectou as relações entre os sectores ambientais constituídos e grupos de variáveis, que se intensificam para ambientes mais "extremos", de maior atlanticidade ou maior continentalidade. A análise pela teoria da informação evidenciou a maior importância das variáveis que definem os climas atlânticos, tanto para o conjunto da informação total mútua, como para a elaboração da classificação do "TWINSPAN". A comparação dos resultados obtidos com as zonagens publicadas anteriormente evidenciam o carácter específico de cada uma delas: as Grandes Regiões (Agroconsultores e COBA, 1990) valorizam os factores fisiográficos e de uso do solo, conferindo-lhe um carácter zonal; a Carta Ecológica (Albuquerque, 1945) apresenta uma maior sectorização do território, resultado do cruzamento directo de factores climáticos e fisiográficos; a Carta de Agrotipos (Albuquerque, 1945), pela sua especificidade, bem como pela especificidade do território, é a que mais se aproxima dos resultados do estudo, tanto globalmente como pela importância dada aos factores ambientais. A técnica de amostragem utilizada mostrou-se eficiente, particularmente ao nível da discriminação dos principais ambientes da província de Trás-os-Montes e Alto Douro. Os resultados supõem um

melhoramento da discriminação realizada pelas anteriores classificações, apresentando sempre classes que a algum nível da hierarquia de classificação, incorporam a informação contida nas outras sectorizações.

RESUMÉN

El presente estudio se basa en la utilización de métodos multivariantes y de la teoría de la información para el estudio de la ecología de "Trás-os-Montes e Alto Douro", provincia de Portugal. En la primera parte se plantea la discusión de las principales metodologías multivariantes desarrolladas hasta hoy, en la segunda se prueban las posibilidades de aplicar una técnica simple de muestreo sistemático de las variables ambientales en una escala territorial amplia. Los resultados obtenidos se han confrontado con las sectorizaciones existentes de carácter sintético, regionales o nacionales referidas al territorio de Trás-os-Montes e Alto Douro. Las variables ambientales del tipo climático, fisiográfico y de usos del suelo se obtuvieron de la bibliografía publicada y cartografiada sobre el tema, con una referencia en malla regular cuyos puntos se situaron cada minuto de latitud, a lo largo de meridianos distanciados 15 minutos de longitud. Se han realizado análisis de clasificación, de concentración, y de información. La técnica "TWINSPAN" de clasificación separa claramente los ambientes de influencia atlántica, de los más continentales, ya que las variables climáticas de características atlánticas asumen gran importancia discriminante. Los factores fisiográficos y de uso de suelo ganan importancia en la diferenciación más detallada de los ocho sectores establecidos en el estudio. Entre los sectores de carácter atlántico, se individualizan los ambientes de las cumbres graníticas de utilización forestal y pastoril, las comunidades rurales de montaña con su paisaje agrario de autoabastecimiento, la media montaña con base de esquistos y usos de secano, y las zonas de ribera, de amplias terrazas con cultivos de regadío. En los sectores de menor influencia atlántica, destaca el paisaje vitícola del Duero, sus afluentes con llanos aluviales de cereales y frutales, el altiplano en el que se intercalan zonas secas de cereales con pastizales de secano, y las elevaciones graníticas, sin utilización agraria importante aparte de pequeños espacios de vid y pinar. El análisis de concentración, detectó las relaciones entre grupos de variables y los sectores ambientales definidos, cuya importancia se intensifica para los ambientes más "extremos", ya sea de tipo atlántico o de tipo continental. La teoría de la información clarifica la gran importancia de las variables que definen el clima atlántico, tanto al nivel del aporte de información compartida total, como en la composición de la clasificación presentada. La comparación entre los resultados obtenidos con las anteriores clasificaciones identifica la especificidad de cada una de ellas: las "Grandes Regiones" (Agroconsultores e COBA, 1990), valora la fisiografía y los tipos de uso de suelo, lo que le confiere gran carácter zonal; la "Carta Ecológica" (Albuquerque, 1945) presenta una gran sectorización del territorio, con gran número de clases, derivado de sobreponer

directamente las condiciones climáticas con las de suelo; su especificidad hace que la "Carta de Agrotipos" (Albuquerque, 1945) sea la que mejor se acerca a los resultados de este estudio, tanto por su enfoque, como por la metodología empleada. La técnica de muestreo planteada en este trabajo se ha mostrado eficiente, sobre todo para un reconocimiento rápido que permita una discriminación suficiente de los distintos ambientes de "Trás-os-Montes e Alto Douro". Los resultados suponen una mejora en la discriminación que aportaban las clasificaciones anteriores y en todos los casos presentan clases que a algún nivel de la jerarquía de clasificación, recogen la información que aportaban las sectorizaciones anteriores.

RÉSUMÉ

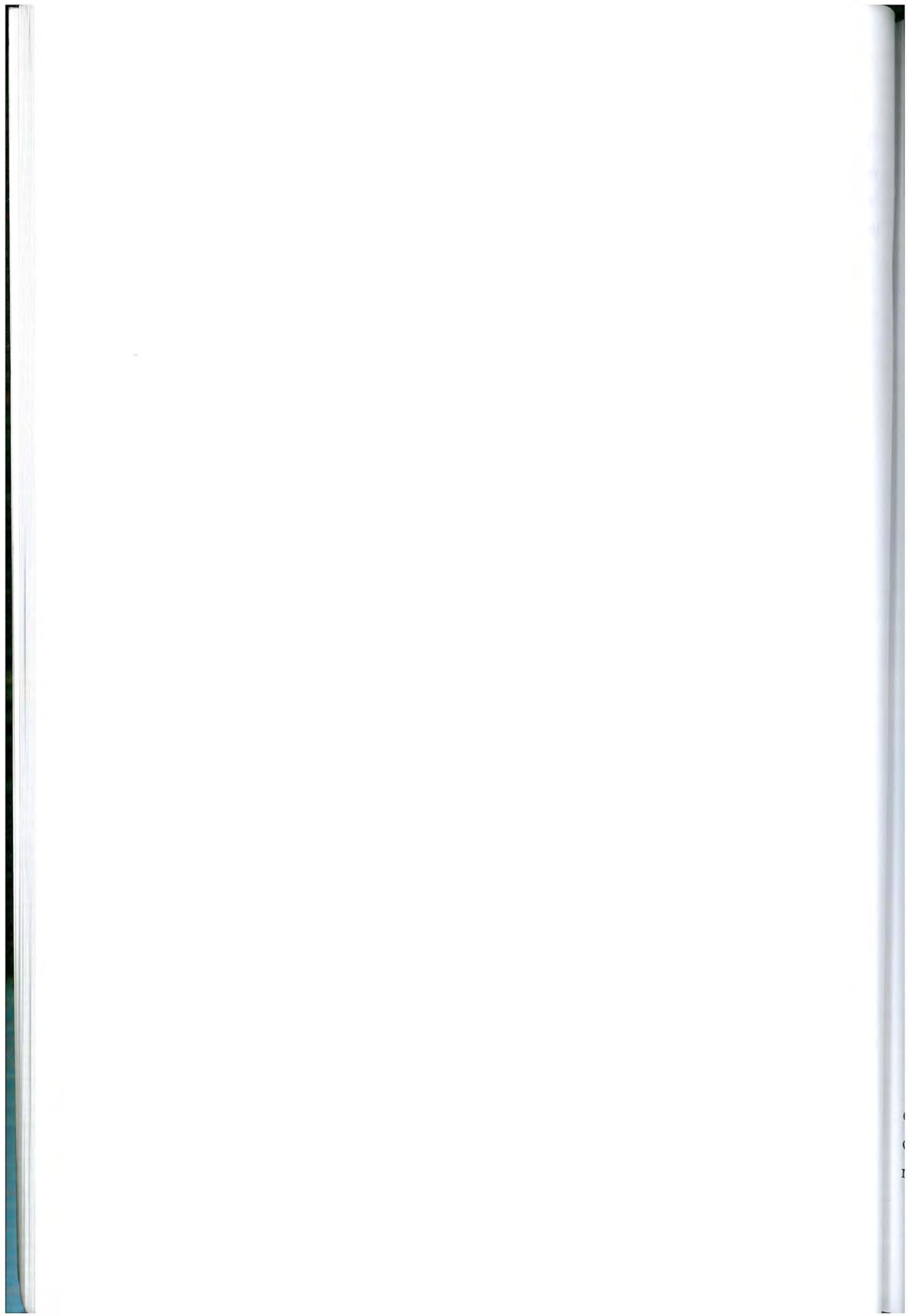
Cet étude utilise des méthodes multivariées et de la théorie de l'information pour l'étude de l'écologie de "Trás-os-Montes e Alto Douro". Pour cela, on discute les principales méthodologies multivariées qui ont été développées jusqu'au présent, de même que l'on teste une technique simple d'échantillonnage systématique des principales variables de l'environnement de cette province. Les résultats ont été confrontés avec d'autres sectorisations écologiques, soit régionales, soit nationales. L'information a été prise dans diverse bibliographie cartographique publiée sur ce thème, au moyen d'un échantillonnage en grille sur chaque minute de latitude, en méridiens éloignés 15 minutes de longitude. On a réalisé des analyses de classification, de concentration et d'information. On a utilisé la technique classificatrice du "TWINSPAN", qui a séparé les sites avec influence atlantique, de ceux d'influence continentale, opération pour laquelle a tenu importance les facteurs climatiques du type océanique. Les facteurs du type physiographique et de l'usage du sol sont importants quand on particularise les huit secteurs établis dans cet étude. Dans les secteurs atlantiques, on a individualisé les ambiances des sommets granitiques d'utilisation forestière et pastorale, les ambiances des agglomérations rurales avec leur paysage d'autoconsommation, la montagne moyenne d'origine schisteuse et utilisation agricole non-arrosée, et les zones de rivière, avec de larges terrasses et cultures des terrains arrosables. Dans les secteurs de moindre influence atlantique, il y a le paysage viticole du "Douro", les vallées de ses affluents avec utilisation par des céréales et arbres de fruits, le plateau dominé par des céréales avec quelques pâturages, et les élévations granitiques sans utilisation agricole principale, mais avec quelques espaces de vigne et pins. L'analyse de concentration a mis en évidence les relations entre les secteurs constitués et les groupes de variables, plus remarquables dans les extrêmes, soit atlantiques soit continentaux. L'analyse par la théorie de l'information a montré l'importance des variables climatiques atlantiques, pour l'ensemble de l'information mutuelle totale et pour l'élaboration de la partition de "TWINSPAN". La comparaison des résultats obtenus avec les zonages publiés a montré le caractère spécifique de chacun d'eux: les Grandes Régions (Agroconsultores e COBA, 1990) valorisent les facteurs de

physiographie et d'usage du sol; la Carte Ecologique (Albuquerque, 1945) présente une plus grande sectorisation du territoire, en résultat de l'intersection des facteurs climatiques et physiographiques; la Carte de Agrotypes (Albuquerque, 1945), pour sa spécificité et pour la spécificité du territoire, est celle qui ressemble le plus aux résultats de l'étude, tant globalement, que pour l'importance apportée par les facteurs. La technique d'échantillonnage a été très efficace, particulièrement au niveaux de la discrimination des principales ambiances de la province de "Trás-os-Montes e Alto Douro". Ses résultats montrent que les classes définies par "TWINSPAN" sont les classes plus abondantes, aussi bien que les écologies plus "extrêmes" des partitions qui sont publiées.

ABSTRACT

This study presents an application of multivariate methods and information theory analysis in studying the ecology of "Trás-os-Montes e Alto Douro". So, the most important multivariate methodologies developed until today are discussed, and a simple systematic sampling method of the environmental variables in that region of Portugal tested. Results of a clustering technique are confronted with other ecological classifications of national and regional level. The information was taken from various cartographic bibliography published about the theme, for samples localised at every minute of latitude, over meridians distanced fifteen minutes of longitude. A classification, concentration and information analysis are undertaken. The clustering technique "TWINSPAN" separated the sites of oceanic influence from those of continental influence, with the exclusive participation of oceanic climatic factors. Physiography and land use factors are important in the particularising of the eight sectors established. The oceanic sectors are divided in the granitic mountainous summits with a forest and pasture use, the rural agglomerates with a subsistence landscape, the gentle schist slopes with dry land cultures and the low wet zones organised in large agrarian terraces. The four continental sectors individualise the viticultural landscape of "Douro", the intermediate valleys of their affluents with dry land cereals, granitic hills with some forest and vineyards and the great schist plateau of dry cereals and wetter pasture areas. The concentration analysis detected the relationship between the established sectors and groups of variables, remarkable at the "extremes" ecologies, more or less oceanic. The information analysis shows the important rule of oceanic climatic variables to the mutual information total and in the groups made by "TWINSPAN". The specific characters of the previous empiric zonations are set in evidence when they are compared with the results of "TWINSPAN": in Great Regions (Agroconsultores e COBA, 1990), physiography and land use variables are more important than in others classifications; in the Ecological Map (Albuquerque, 1945), the excessive number of sectors results from a direct overlap of types of soil and climate; the Agrotypes Map (Albuquerque, 1945), with your own specificity and the specificity of the territory of "Trás-os-Montes e Alto Douro", is the more similar to the

"TWINSpan" clustering, in a global sense and according to the meaning of the environmental variables. The sampling technique proved effective, particularly in the discrimination of the "environments" of that region. The results show that clusters of "TWINSpan" contain the more important and more specific category of the others zonations published so far.



CAPÍTULO 1

INTRODUÇÃO

A Província de Trás-os-Montes e Alto Douro corresponde ao extremo Nordeste de Portugal e tem uma superfície de aproximadamente 21 600 Km². O seu território abrange a totalidade dos distritos de Bragança e Vila Real e parte dos distritos de Viseu e Guarda. Tem uma forma aproximadamente rectangular, limitada entre os meridianos 8°W e 6°W e os paralelos 41°N e 42°N da Rede Geodésica Internacional.

As três forças geográficas definidas por Ribeiro e Lautensach (1987) como indutoras dos diferentes gradientes da Península Ibérica cruzam-se no território de Trás-os-Montes e Alto Douro: as "planetárias" com sentido Norte-Sul, as "continentais" com sentido Nordeste-Sudoeste e as "atlante-mediterrâneas" com sentido Noroeste-Sudeste. As "planetárias" induzem a transição das paisagens húmidas da Europa do Norte para as norte-africanas; as "continentais" representam a variação do litoral para o interior que se faz sentir sobre os vários aspectos geográficos; as terceiras representam o confronto entre um clima de carácter atlântico (influência oceânica) e outro do tipo mediterrâneo (verão seco e quente). De tal confronto resultam as diferentes áreas climáticas observados na província; a um gradual esbatimento das influências oceânicas de Oeste para Este, contrapõe-se a acentuação das características climáticas mediterrânicas com sentido Norte-Sul.

A conformação actual da Província baseia-se na superfície aplanada do maciço Hespérico de constituição xistosa, interrompida por depressões e elevações. As depressões mais importantes são o vale do rio Douro e os seus afluentes. O vale do Douro contorna o lado oriental da Província, inflectindo para Oeste, atravessando o sul do seu território. As elevações correspondem geralmente às montanhas residuais da erosão do Massiço Hespérico, de litologia mais resistente, às quais se juntam a Ocidente as cadeias montanhosas do designado Eixo de Culminação Ibérica Principal. Este eixo tem origem nas diversas fases de enrugamento devidas à orogenia hercínica e manifesta todo o seu esplendor nas serras de Montemuro, Marão e Alvão. É o designado sistema montanhoso Galaico-Duriense, que isola a região oriental da província da região mais próxima do litoral, fazendo desaparecer por completo a pene-planície (Ribeiro e Lautensach, 1987).

Ponto de encontro entre os domínios biogeográficos Mediterrâneo e Eurosiberiano, o seu clima é profundamente marcado pela fisiografia do território, ao qual ainda se sobrepõe influências oceânicas das regiões mais ocidentais. Daí resulta então um eixo longitudinal de clima mediterrânico, constituído pelo vale do Douro e seus afluentes. O regime térmico caracteriza-se por

Invernos doces e Verões sufocantes, propriedades que podem acentuar-se nos vales interiores da província (Tua e Sabor). O regime pluviométrico é seco com acentuada sazonalidade. Nas regiões de montanha, os regimes invertem as suas tendências, caracterizando-se os Invernos por serem rigorosos e húmidos e os Verões suaves. A região planáltica caracteriza-se por Invernos rigorosos e Verões sufocantes, com um regime pluvial seco.

Sobre a base física da província, o Homem foi estabelecendo os seus modelos de utilização do espaço, dos quais se destaca pela sua espectacularidade, a organização em socalcos do Douro Vinhateiro. Grande parte do território é utilizado para o cultivo de cereais acompanhado de culturas mediterrânicas perenes como a oliveira (*Olea europaea*), a amendoeira (*Prunus amygdalus*), e o castanheiro (*Castanea sativa*), intervalado por regiões de montanha, ocupadas por matos e floresta nas superfícies côncavas, e prados permanentes nas superfícies convexas. Da vegetação original subsistem hoje magníficas florestas de carvalho (*Quercus pyrenaica*), sobreiro (*Q. suber*) e azinheira (*Q. rotundifolia*).

A oportunidade da realização deste trabalho deve-se, antes de mais, à presente "revolução" da utilização do espaço induzida por dois factores importantes que influenciam a utilização futura do território da Província. Em primeiro lugar, o fenómeno de desertificação das comunidades rurais pelas mais diversas contingências da evolução da sociedade actual; como resultado, tem-se assistido a um envelhecimento acentuado dos utilizadores directos desse espaço, sendo abandonado quando esses mesmos perdem a sua capacidade de intervenção. O segundo são as alterações induzidas nestes sistemas agrários pela ampliação das suas zonas de influência, ampliação essa devida à criação de infra-estruturas de comunicação e da abolição de fronteiras comerciais. Este segundo factor, mais recente que o anterior, manifesta-se actualmente pela alteração abrupta dos custos dos factores de produção e das suas consequências sobre o rendimento da actividade agrária.

O estudo das implicações desta revolução, bem como das possíveis respostas aos seus efeitos, deverá ser feito de uma forma tipificada, tendo em linha de conta toda a estrutura e funcionamento equilibrados da paisagem afectada. Este trabalho pretende contribuir para a criação duma base ecológica e de paisagem, consubstanciada numa classificação conseguida por um tipo de amostragem expedito, que sustente as decisões em termos de racionalização do uso do território e se constitua como ponto de partida para a elaboração de soluções alternativas aos sistemas de produção actuais, de acordo com as características do espaço em que esses sistemas se integram.

A organização deste texto apoia-se em duas grandes partes: um capítulo onde são revistas as principais metodologias de análise multivariada e análise de informação aplicadas à ecologia, e outro capítulo onde é aplicada uma técnica de classificação a uma amostragem simples e expedita

da estrutura biofísica da província de Trás-os-Montes e Alto Douro, bem como técnicas de avaliação da classificação obtida.

A panóplia de metodologias de análise de factores ecológicos existentes hoje em dia, justificam a sua revisão e organização num capítulo próprio; são técnicas que disponibilizam ferramentas "potentes", que permitem explicar e visualizar a estrutura e funcionamento de gradientes ambientais. Destacam-se as técnicas de análise multivariada, classificação e ordenação, e a teoria da informação. As primeiras permitem o delineamento dos gradientes, base conceptual de entendimento das variações ecológicas que se sucedem ao longo do território; quando aplicadas a estudos de índole espacial, as técnicas de análise multivariada permitem dividir o território em zonas consideradas homogéneas em função dos factores estudados e nível de detalhe utilizado, bem como indicar a posição relativa dos factores ecológicos em jogo. A teoria da informação, quando aplicada ao reconhecimento dos sectores definidos, avalia a sua ordem interna e compara-a com a globalidade do território, e aponta ainda "indicadores" cujo papel na organização do território será fundamental à manutenção da ordem estabelecida.

A diversidade de complexos de factores ecológicos presentes no território transmontano justifica a sua abordagem sistemática, de forma a clarificar a sua estrutura, bem como a comparar os resultados com outras classificações climáticas, pedológicas e de vegetação, baseada na observação empírica, quer de âmbito nacional, quer de âmbito regional; de entre estas destacam-se a Carta Ecológica e a Carta de Agrotipos de meados do século (Albuquerque, 1945; Albuquerque, 1954), e mais recentemente a Carta dos Solos, Carta de Utilização Actual do Solo e Carta de Uso Potencial do Solo do Nordeste de Portugal (Agroconsultores e COBA, 1990). A classificação do território é uma das aplicações práticas mais interessantes da Ecologia de Paisagem ("Landscape Ecology"). Segundo VINK (1983), "Land classification basically means the arrangement of land in class". Sendo, em termos gerais, mais do que uma avaliação do território, a classificação tem por finalidade a integração dos conhecimentos (factores, elementos, variáveis, etc.) que concorrem para a identificação de determinado espaço. Têm sido vários os modelos concebidos para a realização dessa integração. Na elaboração das técnicas de classificação nota-se um esforço por parte dos seus autores para a concepção de uma metodologia holística, que lhe permita ser a base de qualquer estudo de ordenamento do território. No entanto, na sua aplicação prática constata-se sempre a necessidade da sua adaptação ao caso concreto de cada estudo, que varia naturalmente com a escala, as actividades em alternativa, a ecologia do território, etc.

Os objectivos finais do trabalho apresentado repartem-se pela discussão e utilização de metodologias de análise multivariada e de informação, pela experimentação de uma técnica sistemática de inventário do meio com base num referencial linear notável, e pela elaboração de uma zonagem analítica do território, seguindo a ordem natural dos factores que regem o seu

comportamento. Tais objectivos estão interligados e sucedem-se de forma a confirmarem-se e apoiarem-se entre si.

O primeiro destes objectivos pretende enquadrar as análises aplicadas ao território transmontano dentro das modernas e complexas abordagens de dados ecológicos. Tenta-se desta forma minimamente exaustiva compilar o contributo que os vários autores têm produzido e descrito sob as mais diversas ópticas, ao longo dos últimos anos. A forma de apresentação pretende unificar a abordagem das técnicas de análise multivariada e de análise de informação, por forma a entender-se o papel que cumpre a cada método; isso implica uma descrição simplificada que sacrifica a profundidade da sua discussão à sua apresentação mais completa.

O segundo destes objectivos testa a eficácia de um método expedito de recolha de informação sobre o território em estudo, de uma forma ajustada ao seu tratamento segundo as técnicas descritas anteriormente, bem como à possibilidade de comparação com classificações específicas realizadas antes. A sua sistematização por meridianos da Rede Geodésica Internacional permite a utilização de toda a informação cartografada na obtenção dos resultados finais, bem como a eliminação da subjectividade que as classificações empíricas acarretam. Deste modo, integra-se informação dispersa por diferentes cartografias temáticas (clima, topografia, solos, etc.), num nível que permite identificar os gradientes mais importantes da província, em equilíbrio com o rigor dos resultados e a operacionalidade dos mesmos.

O terceiro objectivo pretende tirar partido da análise multivariada e teoria da informação, utilizando-as para reconhecer, identificar e caracterizar os sectores ecológicos mais importantes do território da província de Trás-os-Montes e Alto Douro, nos quais se fundamenta toda a utilização produtiva do seu espaço. Tais sectores são avaliados por "indicadores" próprios, os quais representam a base conceptual do entendimento do seu território. Os objectivos são, antes de mais, a identificação de áreas de comportamento semelhante quanto à sua situação actual, bem como quanto às respostas possíveis a alterações das condições exógenas que regulam a sua utilização pelo Homem. Os resultados obtidos são comparados com as classificações já existentes, identificando os factores de maior responsabilidade na sua construção.

As diversas áreas homogéneas identificadas deverão possuir comportamentos típicos de resposta às novas condições que acima foram descritas, apresentando áreas "marginais" semelhantes, terrenos abandonados do mesmo tipo, tanto em termos físicos como de vegetação colonizadora (pioneira), e a mesma história de utilização anterior. As utilizações alternativas deverão preservar o equilíbrio do território em termos estruturais e funcionais, com a sua produtividade adequada às suas potencialidades. Deste modo, tornar-se-á possível um estudo

dirigido e aprofundado da problemática de cada zona, sem se cair em soluções redundantes ou em abordagens pontuais que depois não encontram a devida expressão no território.

l
)
>
a

,
o
.o
is
a
to
lo
os

de
is"
ção
vas
sua
ido

CAPÍTULO 2

REVISÃO DE METODOLOGIAS PARA ANÁLISE DE DADOS

2.1 ANÁLISE DE COMPLEXOS DE DADOS ECOLÓGICOS

A análise de dados em ecologia serve-se hoje de disciplinas que nem sempre foram comuns às ciências naturais. Para isso contribui a necessidade de compreender a complexidade dos sistemas naturais, revelada pelo avanço do conhecimento ao longo dos tempos. Hoje em dia, a avaliação de qualquer componente, fluxo ou relação caracterizadores de um sistema, recorre a técnicas numéricas anteriormente exclusivas da física, matemática, sociologia, etc. De entre essas, Legendre e Legendre (1984) destacam o cálculo matricial, a análise dimensional e multidimensional e a análise pela teoria da informação (figura 2.1).

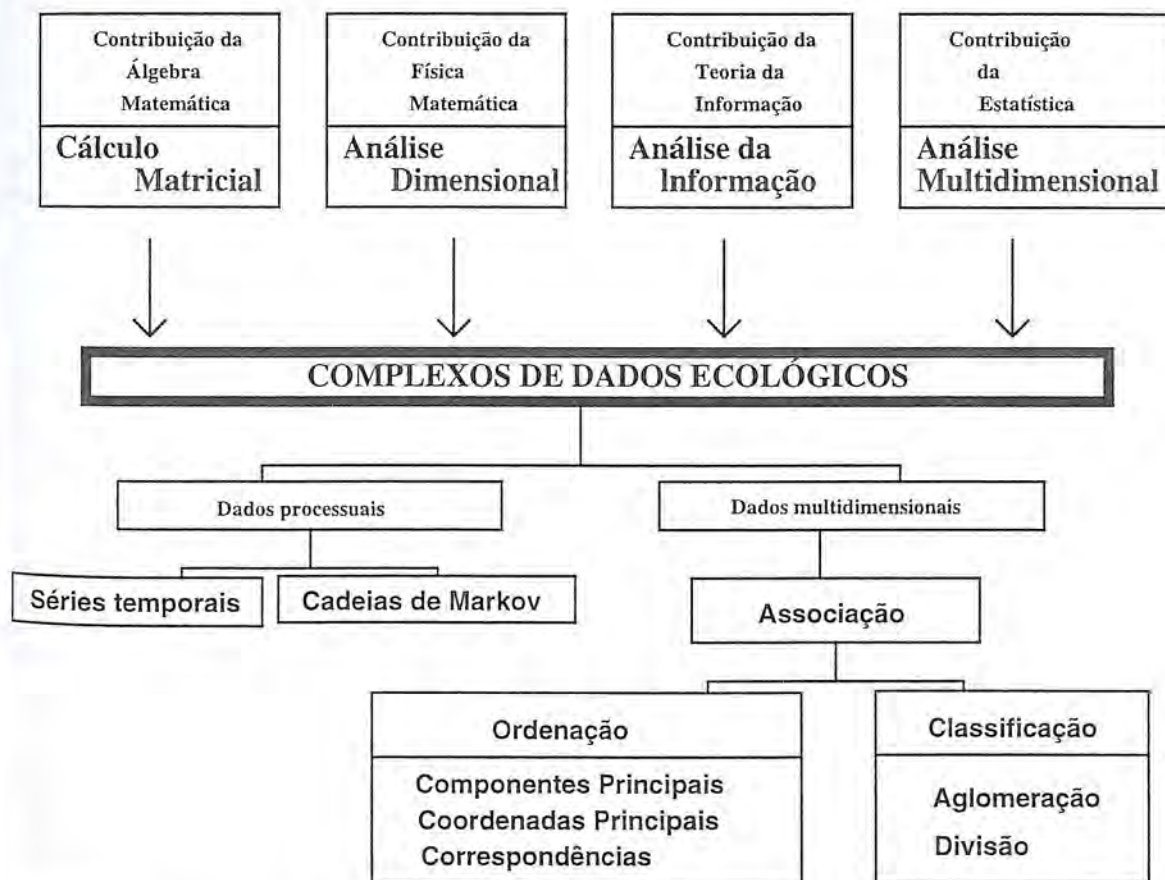


Figura 2.1- Síntese das contribuições de outras ciências à investigação em ciências do meio natural (adapt. Legendre e Legendre, 1984a)

O cálculo matricial é uma forma cómoda e natural de trabalhar os dados em ecologia. A álgebra matemática tem a vantagem de ser simples, eficaz, familiar aos cientistas, e ainda adequada ao tratamento de dados em computador; a física matemática, nomeadamente a análise de dimensões, permite, senão resolver, pelo menos comprovar a correcção das análises realizadas; a teoria da informação conduz à generalização do tratamento de dados, que de outra forma não são fáceis de conseguir e demonstrar; os métodos estatísticos proporcionam a análise multidimensional (Legendre e Legendre, 1984).

Os dados obtidos na investigação de qualquer sistema natural devem reflectir, naturalmente, a complexidade inerente ao objecto em estudo. Em ecologia, esses complexos de dados dividem-se em *processuais* e *multidimensionais*, consoante se considere ou não o factor tempo. Os complexos de dados *processuais* reflectem a evolução dum sistema, quer no espaço, quer no tempo. São exemplos de dados desse tipo os derivados da investigação de sucessões, séries de vegetação, etc., que para isso utilizam metodologias como as séries temporais e as cadeias de Markov. Os complexos *multidimensionais* reflectem a natureza dos sistemas que descrevem; resultam de um sem número de causas, o que influencia de uma forma mais ou menos marcada os seus resultados (Legendre e Legendre, 1984).

Segundo Gauch (1982), em ecologia de comunidades ou de paisagem, a entidade abordada é caracterizada pelo seu grande número de atributos redundantes com relações internas, complexos, e grosseiros pela sua difícil avaliação. Muitas vezes, a informação apenas é interpretada indirectamente pelas propriedades emergentes do carácter integrador dos objectos em estudo em ecologia.

Cattell (1966; citado por Legendre e Legendre, 1984), apresenta uma visão original e interessante dos complexos de dados ecológicos, abordando as análises de dados ecológicos como operativas sobre um bloco tridimensional, onde o tempo aparece como uma dimensão da diversidade dos dados. Para a variação de dados ecológicos concorrem diferentes objectos (indivíduos), características (variáveis) e momentos (tempos) (figura 2.2); a análise das suas relações designar-se-ia por **O** - entre momentos para um conjunto de variáveis, **P** - entre variáveis para um conjunto de momentos, **Q** - entre objectos para um conjunto de variáveis, **R** - entre variáveis para um conjunto de objectos, **S** - entre objectos para um conjunto de momentos e **T** - entre momentos para um conjunto de objectos.

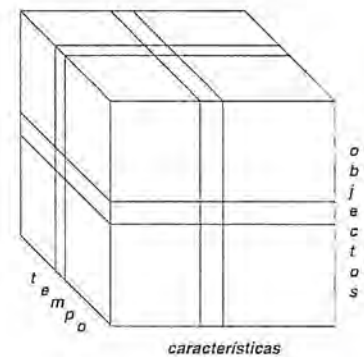
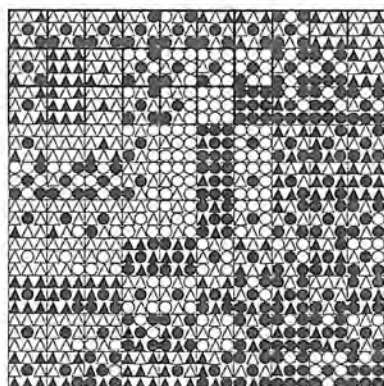


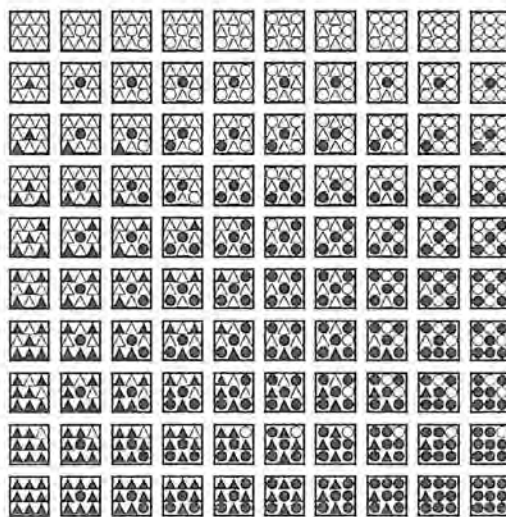
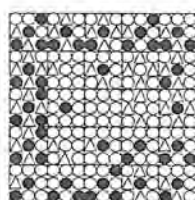
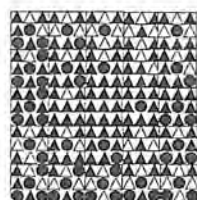
Figura 2.2 - Estruturação conceitual de dados ecológicos (adapt. Legendre e Legendre, 1984a).

2.2 BREVE RESENHA DE MÉTODOS DE ANÁLISE MULTIVARIADA

Embora as primeiras análises com um grande número de variáveis remontem ao início deste século, a grande maioria das suas bases teóricas actuais foram estabelecidas de forma definitiva apenas por volta de 1930, com complementos importantes após 1950 (Lefebvre, 1983). As metodologias empregues em análise multivariada dividem-se em dois grupos: métodos de **classificação** e métodos de **ordenação**. Ambas se baseiam na semelhança entre os indivíduos/objectos, tendo em conta as variáveis/atributos que os caracterizam ou, de uma forma recíproca, no comportamento semelhante das variáveis, nos indivíduos/objectos em que coincidem.



I



II

III

Figura 2.3 - Representação esquemática de um espaço (I), com o respectivo modelo de classificação (II) e ordenação (III), tendo em conta 9 atributos por objecto, com 4 classes cada (Δ , \blacktriangle , \circ e \bullet) (Adapt. Zonneveld, *in press*).

No exemplo da figura 2.3, adaptado de Zonneveld (*in press*), representa-se o que pode ser a disposição real dos objectos/indivíduos de uma paisagem (I) e os respectivos modelos teóricos; as unidades de amostragem são caracterizadas por 9 atributos, descritas pelas suas formas, *triangulares*

ou *circulares*, e pela sua cor, *branca* ou *negra*. No primeiro modelo teórico (II), um modelo de classificação, as unidades de amostragem são agrupadas em quatro classes que se desenvolvem em torno de elementos extremos do conjunto inicial (neste caso *triângulos brancos*, *triângulos negros*, *círculos brancos* e *círculos negros*), com base nas respectivas semelhanças. No segundo modelo teórico (III), um modelo de ordenação, as unidades de amostragem são ordenadas segundo dois gradientes, atendendo às suas características; das formas triangulares para as circulares, e da cor branca para a negra.

Os complexos de dados organizam-se em matrizes \mathbf{X} , nas quais cada valor x_{ij} corresponde ao observado para a variável j (valor, ordem ou frequência) no objecto i :

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \dots & \dots & x_{ij} & \dots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{pmatrix}$$

A natureza das variáveis/atributos que descrevem determinado indivíduo/objecto podem considerar-se dependentes ou independentes entre si. Em muitos estudos, é precisamente essa dependência o objectivo de investigação, sendo nesses casos, a matriz de partida do tipo da apresentada abaixo, onde as m variáveis se dividem em p variáveis independentes e q variáveis dependentes das primeiras.

$$(\mathbf{X} : \mathbf{Y}) = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1p} & y_{11} & y_{12} & \dots & y_{1q} \\ x_{21} & x_{22} & \dots & x_{2p} & y_{21} & y_{22} & \dots & y_{2q} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{np} & y_{n1} & y_{n2} & \dots & y_{nq} \end{pmatrix}$$

Também a natureza dos objectos/indivíduos em estudo pode levar à sua diferenciação, ainda antes do seu tratamento numérico. Nesse caso, interessa conhecer o valor e/ou a significância dessa divisão; os dados apresentam-se em o grupos, cada um com certo número de objectos/indivíduos (n, \dots, p), descritos por m variáveis.

$$(\mathbf{X}) = \begin{pmatrix} x_{111} & x_{112} & \dots & x_{11m} \\ \dots & \dots & \dots & \dots \\ x_{1n1} & x_{1n2} & \dots & x_{1nm} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ x_{o11} & x_{o12} & \dots & x_{o1m} \\ \dots & \dots & \dots & \dots \\ x_{op1} & x_{op2} & \dots & x_{opm} \end{pmatrix}$$

2.2.1 MÉTODOS DE CLASSIFICAÇÃO

Zonneveld (*in press*) refere-se ao termo classificação ("clustering" em inglês, "groupement" em francês) como, o "calão científico" utilizado para definir a "arrumação" sistemática dos objectos em estudo. Segundo o mesmo autor, os objectos clássicos de classificação científica são os organismos existentes, sendo a primeira acção do homem no paraíso (Genesis 2:19). Os organismos são indivíduos claramente definidos, exemplos clássicos de complexos individualizados. Legendre e Legendre (1984) definem o agrupamento de objectos como, para além do reconhecimento do seu grau de semelhança, a percepção da diferença ou afastamento entre os grupos constituídos. Se para o taxonomista, que conhece as descontinuidades verificadas ao nível infra-específico, tal tarefa não é tão árdua, as dificuldades são maiores para os ecólogos classificarem o seu objecto de estudo, pela natureza de *continuum* do meio ambiente. A aplicação de técnicas da taxonomia directamente à ecologia, sem a necessária reflexão teórica, é a razão de muitos e decepcionantes resultados obtidos. Classificação é uma forma explícita de identificação de grupos de objectos/indivíduos, ajudando à compreensão da estrutura da variação das suas características. Para uma estrutura contínua, a análise de classificação impõe um arranjo agrupado, um *continuum* que é separado num sistema descontínuo de tipos ou classes.

Hoje em dia existem modelos que classificam atributos do território relativos ao solo e à vegetação. Segundo Lefebvre (1983), os seus objectivos são:

- a) informar sobre a ocorrência conjunta dos indivíduos (estrutura interna dos dados);
- b) estabelecer associações típicas para estudos descritivos (sintaxonomia e cartografia);
- c) discriminar essas associações em relação ao seu ambiente, pela análise dos grupos formados e respectivas variáveis ambientais caracterizadoras (análise externa).

As técnicas de classificação são mais recentes que as de ordenação, e desenvolveram-se sobretudo a partir da massificação do uso dos computadores. Recorrem, tal como as de ordenação, a matrizes simétricas de associação como ponto de partida do seu desenvolvimento. As técnicas usadas em ecologia são numerosas, tornando necessário recorrer à sua sistematização. As separações mais usuais são entre técnicas *aglomerativas/divisivas* e entre *hierárquicas/não hierárquicas*. As técnicas aglomerativas, ou classificação ascendente, conduzem a classes com carácter diferente das divisivas, o que aconselha a serem utilizadas de uma forma integrada; partem de um baixo nível hierárquico, agrupando unidades básicas de acordo com as suas semelhanças. É também denominada **Tipificação**. As técnicas divisivas, ou classificação descendente, partem de uma alta hierarquia e dividem a totalidade dos objectos em segmentos, baseados nas suas diferenças (Zonneveld, *in press*).

As técnicas *hierárquicas* distinguem-se das *não hierárquicas* por a classificação final se atingir em etapas sucessivas, às quais correspondem classificações intermédias. O utilizador tem assim uma ideia da evolução da aglomeração ou divisão dos objectos que pretende classificar. No caso de técnicas aglomerativas, em cada uma das suas etapas, os objectos ou classes de ordem inferior são agrupadas em classes de ordem superior. No caso das técnicas divisivas, o todo inicial, constituído pela totalidade dos objectos ou classes de si derivadas, é sucessivamente segmentado em classes de ordem inferior, correspondendo cada etapa a uma determinada classificação.

As técnicas de classificação são também designadas por *monotéticas* ou *politéticas*; as *monotéticas* utilizam apenas a informação de um atributo para decidir unir ou separar classes, para o qual deverá ser o que melhor caracteriza cada um dos objectos, ou seja, com mais informação comum a todos os que são utilizados na caracterização dos objectos. Nas *politéticas*, são tidos em consideração vários atributos para essa tarefa. O tipo de algoritmo de classificação permite distinguir as classificações em *simultâneas* e *sequenciais*, atendendo ao número de etapas em que é atingido o resultado final. Nas primeiras, o resultado final é atingido em uma só etapa, baseando-se geralmente em técnicas de ordenação. As *sequenciais* atingem o resultado final depois de uma série de etapas recorrentes.

À partida, o número de grupos utilizados pela classificação não deve ser imposto; Podani (1990) recomenda que se realize primeiro uma classificação do tipo hierárquico, após a qual se reconheça a mais natural partição dos dados. Os resultados podem ser melhorados por outra classificação posterior, não hierárquica, na qual se força a constituição desse número determinado de classes. Grande parte dos métodos de classificação são fruto do desenvolvimento de trabalhos de Lance e Williams (1967; citados por Jongman e col, 1987). Alguns deles foram seleccionados para serem descritos neste trabalho; a sua escolha advém do facto de serem clássicos, usuais no tratamento de dados de sistemas ecológicos e/ou comuns nos programas informáticos de estatística.

2.2.1.1 Métodos hierárquicos aglomerativos

Podani (1990) descreve o que considera como sendo as quatro formas principais de realizar uma classificação hierárquica de modo aglomerativo: *optimização da distância*, *optimização de homogeneização*, *utilização de critérios da teoria da informação* e *optimização global*, as quais se descrevem brevemente em continuação.

OPTIMIZAÇÃO DA DISTÂNCIA

É o método de classificação mais clássico, e o seu algoritmo resume-se à ordenação decrescente dos pares de objectos/indivíduos, em função de um qualquer índice de semelhança, e à sua associação em grupos ou com grupos, obedecendo à ordem estabelecida antes. Estas etapas são

repetidas até se atingir um só grupo, recalculando-se entre cada uma das fases a semelhança entre os indivíduos e/ou grupos recém-formados pelo índice antes escolhido. A distância entre grupos (diferença nas suas características) pode ser avaliada por um dos 8 métodos de avaliação da distância descritos por aquele autor (Podani, 1990), o que, combinado com cerca de 30 diferentes possibilidades de calcular a matriz de semelhança, resulta em mais de duas centenas de diferentes possibilidades de obtenção de uma classificação por esta técnica. Na figura 2.4 exemplifica-se como um objecto seria atribuído a três grupos distintos, caso fosse considerada a distância *simplex*, *completa* ou *ao centróide*.

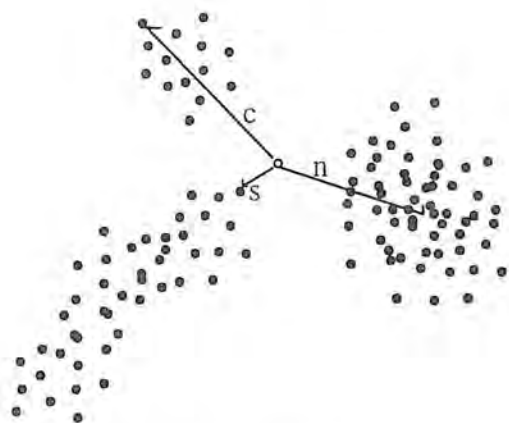


Figura 2.4 - Representação das distâncias de decisão para atribuição do objecto "o" a três grupos diferentes (vizinho mais próximo, centróide e elemento mais afastado).

OPTIMIZAÇÃO DA HOMOGENEIDADE

Neste segundo método, o conceito de "distância" é substituído pelo de "homogeneidade", procurando-se a sua maximização dentro dos grupos formados, em lugar da maximização das "distâncias" entre estes. Podani (1990) propõe três medidas para a homogeneidade (erro da soma dos quadrados, variância e distância média dentro do grupo), o que multiplicado por dois critérios de fusão (minimização do decréscimo da homogeneidade e maximização da homogeneidade dentro dos grupos formados) permite dispôr de 6 diferentes formas de obter a classificação. Podem ser utilizadas cerca de 30 formas diferentes de calcular a matriz de semelhança, bem como recorrer à transformação dos dados iniciais.

CRITÉRIO DA TEORIA DA INFORMAÇÃO

Este método utiliza a teoria da informação apresentada pela primeira vez por Shannon (1948; citado por Orlóci, 1991a). Destina-se a variáveis do tipo binário e apoia-se nos conceitos de *Entropia Diferenciadora* ("pooled entropy") $H(L)$ e *Informação Mútua* $H(I)$, desenvolvidos por Willians e col, (1966); Orlóci (1969); Podani (1980; citado por Podani, 1990). Qualquer dos dois conceitos pode ser utilizado em dois algoritmos - minimização do incremento da heterogeneidade ou minimização dela própria nos novos grupos formados - o que perfaz quatro métodos de realizar

esta classificação. Para além destes algoritmos, outros dois são disponibilizados para realizar a fusão dos grupos: apenas se juntam os vizinhos mais próximos que o são reciprocamente, ou, de uma só vez, os pares mais próximos.

OPTIMIZAÇÃO GLOBAL

O método da optimização global difere radicalmente dos anteriores. A fusão de dois grupos ou objectos depende da qualidade da classificação obtida no seu todo, utilizando como critério avaliador a razão entre a média das "distâncias" dentro dos grupos formados e a média das "distâncias" entre os grupos. Por este método, em cada fase da classificação, dois objectos são agrupados apenas se produzirem o mínimo aumento possível desta razão. Para matriz de semelhança podem ser utilizadas qualquer uma das referidas para os métodos anteriores.

São numerosos os autores que utilizaram e avaliaram as potencialidades destes métodos de classificação. Willians e col (1966) abordaram, sob o ponto de vista conceptual, os métodos aglomerativos hierárquicos e compararam dois métodos de fusão de grupos e cinco índices de semelhança entre objectos. A sua avaliação foi realizada com dados de inventários fitossociológicos, seguindo critérios de perfeita distinção de grupos (bem diferenciados a níveis hierárquicos bem definidos) e significância ecológica, quando comparados com outras informações exteriores (outras classificações, analíticas ou empíricas). Os melhores resultados são obtidos quando utilizam a distância ao centróide como método de fusão de grupos, e a entropia do hipotético novo grupo como critério de semelhança entre grupos ou objectos. Gauch e Whittaker (1981) também compararam três métodos aglomerativos e dois divisivos, concluindo que entre os primeiros, as diferenças não são significativas nem consistentes, quando comparados com os resultados obtidos com os segundos. Para os autores, o facto de serem técnicas que se desenvolvem agrupando pares de objectos, e em que, portanto, os agrupamentos iniciais condicionam os resultados finais, quando utilizados para objectos muitos parecidos ou dissemelhanças inicialmente muito reduzidas, a natureza aleatória dos dados e a intensidade de amostragem poderão condicionar o resultado final. Podani (1989) estabelece e avalia metodologias que permitem comparar os diferentes algoritmos de classificação, recorrendo a uma extensa e importante recolha bibliográfica.

Alvarez e col (1989) utilizaram métodos aglomerativos hierárquicos no agrupamento de quadrículas de amostragem em função de características geofísicas, resultando uma zonagem das condicionantes ecológicas ao uso do território pelo Homem. Gómez Sal e Muñoz-Yanguas (1990), servindo-se do índice de semelhança de Jaccard e do método de fusão por minimização do acréscimo do erro da soma dos quadrados, agruparam diversas proporções de usos e de actuações humanas em função de variadas funções ecológicas do território, do que resultaram vários quadros paisagísticos. Szaro (1990) utilizou as distâncias euclidianas e as distância médias entre grupos para

relacionar 153 locais de vegetação ripícola do Arizona e Novo México, definindo 28 comunidades arbóreas e arbustivas utilizadas para a cartografia da vegetação da região.

2.2.1.2 Métodos hierárquicos divisivos

Os métodos hierárquicos divisivos, mais exigentes ao nível de recursos de cálculo, tiveram um desenvolvimento posterior aos aglomerativos, acompanhando o avanço das técnicas de computação bem como o desenvolvimento recente da fitossociologia, aos quais foram desde sempre especialmente dedicados. De entre eles destacam-se o método de associação e o "TWINSPAN" de Hill (1979).

ASSOCIAÇÃO

A metodologia de associação é, de entre os métodos hierárquicos divisivos, o método mais comum, tendo sido desenvolvido e massivamente utilizado para a definição de associações fitossociológicas, com base em inventários de vegetação. É uma técnica monotética, ou seja, utiliza um atributo apenas para a divisão dos grupos de objectos. Aplica-se a dados do tipo binário (presença/ausência) e o seu desenvolvimento é descrito por Ludwig e Reynolds (1988). Partindo de tabelas de contingência 2×2 para todos os possíveis pares atributos, são calculados os respectivos valores de χ^2 , somando-se para cada atributo, os valores que apresentem significância (3.84 para 1 grau de liberdade e probabilidade de 5%). Considerando-se como atributo divisor aquele que apresentar o maior valor, os objectos são divididos em função da sua presença ou ausência. O procedimento é repetido para cada um dos subgrupos formados, até os grupos encontrados não possuírem qualquer par objecto/atributo com valor de χ^2 significativo.

A utilização do nível de significância de 3.84 leva por vezes à fragmentação excessiva dos objectos em estudo. Vários autores discutem e apresentam soluções para obviar isso. Ludwig e Reynolds (1988) concordam com o método original, preconizando um juízo crítico dos resultados que defina o nível a partir do qual as divisões deixam de ter sentido. No "package" - SYNTAX IV- (Podani, 1990), a escolha do atributo divisor faz-se segundo um dos seguintes critérios: a maximização da soma da *Informação Mútua*, calculada entre os pares de atributos, ou a maximização da *Entropia Diferenciadora*; não é apresentado qualquer critério para o *terminus* das divisões, sendo estas definidas por limitações de software ou pela existência de mais de um atributo divisor com uma distribuição não coincidente. Ao introduzir conceitos da teoria da informação para a escolha do atributo divisor (Podani, 1979), o autor alertou para a necessidade de se recorrer a diferentes métodos de classificação para confirmar as tendências dos resultados obtidos por algum deles.

"TWINSPAN"

O "TWINSPAN", "Two Way Indicator Species Analysis", de Hill (1979), foi também desenvolvido para aplicações em fitossociologia. Ao contrário do anterior, é uma técnica politética, ou seja, pode utilizar mais do que um atributo divisor. Desenvolvido inicialmente para classificar inventários de vegetação, este método permite o uso de níveis de abundância, atributos ordinais, tal como a escala de Braun-Blanquet. Cada nível de abundância é substituído por uma *pseudo-espécie* (novo atributo), como resultado de uma codificação ou transformação de variáveis do tipo ordinal (abundância) em variáveis do tipo cardinal binário (presença/ausência). Em função de limites estabelecidos pelo utilizador, quanto maior for a abundância de uma espécie num inventário, maior será o número de pseudo-espécies que lhe serão atribuídas. Hill (1979) chama a atenção da confusão permitida pelos termos "Indicator Species Analysis" no nome do "TWINSPAN", devido ao facto da ordenação de indicação ser apenas um indicador e não realmente uma base do método.

O algoritmo do "TWINSPAN" inicia-se com a divisão dos inventários pelo centro de gravidade (centróide) do primeiro eixo ordenado da análise de correspondências (ver ponto 2.2.2.1), definindo-se como *espécies diferenciais*, as que surgem nos inventários de apenas um dos lados dessa dicotomia e, atribuindo-se valores positivos e negativos consoante essas preferências sejam por um ou outro lado da ordenação. É atribuído o valor absoluto de preferência 1 por cada atributo (espécie ou pseudo-espécie) que seja três vezes mais frequente num dos lados da dicotomia em relação ao outro, o que permite uma nova ordenação dos inventários, ao somar os valores de preferência das espécies presentes em cada um. Os valores obtidos para cada inventário são "standartizados", de forma a que o seu máximo absoluto corresponda a 1, obtendo-se uma segunda ordenação com base no cálculo da média dos valores de preferência de cada inventário. Esta segunda ordenação pretende que os resultados não sejam afectados pelo pouco peso das espécies mais raras e/ou inventários com poucas espécies, sendo os seus resultados somados com os da primeira ordenação para dar origem à designada *ordenação refinada*, a qual é então dividida pelo seu centro. Com excepção de alguns inventários extremos, esta ordenação refinada, determina a dicotomia final, e os casos localizados perto do ponto de divisão da ordenação refinada são sujeitos a uma terceira ordenação - a *ordenação de indicação*. Esta ordenação de indicação não tem por objectivo atribuir os casos extremos a um dos lados da dicotomia, mas sim reproduzir a dicotomia sugerida pela ordenação refinada, servindo-se de simples funções discriminantes baseadas em poucas espécies mais preferenciais, para caracterizar as suas tendências. Este procedimento é repetido até ao nível de divisão ou dimensão de grupo desejado pelo utilizador.

Na construção de tabelas ordenadas de atributos por inventários são necessários mais dois passos: a ordenação das divisões por comparação dos dois grupos formados a um qualquer nível com outros dois grupos de dois níveis hierarquicamente superiores, e a classificação das espécies de

acordo com a classificação anterior, baseando-se na fidelidade dessas espécies aos grupos de locais, reflectindo deste modo a classificação dos inventários. Deste modo consegue-se construir uma tabela a partir de ambas as classificações (locais e espécies), de forma a ser estabelecida uma aproximação a uma diagonal positiva (desde o canto superior esquerdo ao inferior direito). Os resultados obtidos aparecem deste modo classificados em grupos ordenados ao longo de um gradiente, e a construção dessa tabela de dupla entrada é baseada em características também comuns a outros programas, sendo essa sua característica a grande responsável pela sua utilização em larga escala em ecologia das comunidades e de paisagem (Vos e Stortelder, 1992; Jongman e col, 1987).

Gauch e Wittaker (1981) propõe uma nova técnica hierárquica divisiva e politética baseada na sucessiva divisão "manual" da ordenação realizada pelo primeiro eixo da análise de correspondência "detrended", fundamentada na experiência do investigador. Segundo os autores, a utilidade das divisões realizadas de uma forma subjectiva reside no facto de se desejar separar regiões naturalmente afastadas, de se pretender incorporar julgamentos sobre anteriores avaliações dos dados, ou de as classificações subjectivas serem suficientes para os resultados requeridos.

2.2.1.3 Métodos não hierárquicos

Estas metodologias de classificação são geralmente utilizadas para a obtenção de um número de grupos de objectos pré-definido pelo seu utilizador. Segundo Podani (1990), os resultados dos métodos hierárquicos ajudam na predefinição dos grupos a considerar, sendo melhorados com a utilização das metodologias não hierárquicas, para as quais são descritas a *optimização global*, a "*c-médias indistintas*" e a *árvore de menor distância*.

OPTIMIZAÇÃO GLOBAL

Este método é análogo ao do mesmo nome descrito anteriormente (métodos hierárquicos aglomerativos), em que cada divisão realizada é avaliada pela razão entre a média das distâncias dentro do grupo e a média das distâncias entre os grupos. De uma forma iterativa é seleccionado o objecto responsável pelo maior decréscimo do valor dessa razão, que é deslocado para um novo grupo. Esta metodologia é utilizada para elaborar uma nova classificação ou melhorar outra pré-determinada por técnica diferente; no primeiro caso considera-se uma partição casual dos objectos como ponto de partida do desenvolvimento do método.

Como desenvolvimento desta metodologia, Podani e Feoli (1991) apresentaram uma técnica de classificação simultânea de objectos e atributos. Partindo de uma classificação (aleatória ou baseada em experiências prévias), esta metodologia permite a realocação iterativa de linhas e/ou colunas de uma tabela de objectos por atributos. As partições de objectos e atributos implicam blocos cuja consistência é avaliada pelo χ^2 (qui-quadrado), $H_{(p,q)}$ (entropia diferenciadora) e $S_{(p,q)}$

(soma dos quadrados), consoante se tratem de dados binários, nominais ou reais. A realocação iterativa de linhas ou colunas permite uma crescente consistência dos blocos da tabela, até ao seu nível máximo. O facto de ser um método iterativo, em que os resultados finais dependem da estrutura de partida, requer um prévio conhecimento dos traços gerais da estrutura dos dados, bem como uma crítica final ao resultado.

"C-MÉDIAS" INDISTINTAS

Este método assenta no princípio de que determinado objecto/indivíduo não pertence em exclusivo a um só grupo mas pelo contrário, tem uma cota parte em vários grupos. Essa relação é avaliada com um valor entre 0 e 1, sendo a soma dos valores totais para um objecto igual a 1. Todas as classificações anteriores, quando entendidas desta forma, consideram, os objectos com o valor 1 para uma e só uma classe. Pelas suas características, este tipo de metodologias assemelham-se às de ordenação, podendo os seus resultados ser apresentados ou utilizados para a localização dos indivíduos num espaço multidimensional, cujos eixos definidores seriam os próprios grupos determinados. A distância entre os centróides dos grupos (média das variáveis para cada grupo) é multiplicada pela ponderação dos seus membros. O utilizador, para além do número de grupos requeridos, define ainda o número de iterações e o coeficiente de "indistinção" a ser usado.

ÁRVORE DE MENOR DISTÂNCIA

Este terceiro método não é um tipo de classificação propriamente dito, antes pelo contrário, os seus resultados podem ser usados para conseguir uma classificação do tipo divisivo. Cada objecto a classificar é representado por um ponto e entre cada dois pontos existirá uma ligação proporcional à distância desses dois objectos. Essa ligação é única e não permite ligações duplas e indirectas entre indivíduos. Passo a passo, as ligações de maior tamanho vão sendo desfeitas para dar lugar a subgrupos de objectos. Após a separação de cada cadeia, as distâncias são de novo calculadas dentro de cada subgrupo. O seu resultado final é igual ao do método aglomerativo hierárquico, no qual são minimizadas as distâncias simples entre objectos (ponto 2.2.1.1), proporcionando no entanto, a informação sobre os pares de objectos responsáveis pela fusão ou divisão dos grupos (não disponibilizada no método aglomerativo).

2.2.2 MÉTODOS DE ORDENAÇÃO

As análises de ordenação têm por objectivo a representação de objectos num espaço de uma ou mais dimensões, cujos eixos geradores são definidos pelos atributos utilizados na caracterização desses objectos. A grande maioria das técnicas de ordenação baseia-se na decomposição ou análise espectral de uma matriz simétrica. Por decomposição entende-se a expressão de uma matriz como produto de duas ou mais matrizes, que no caso particular da decomposição espectral, é:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ii} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{ni} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1i} & \dots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \dots & \gamma_{2i} & \dots & \gamma_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \gamma_{i1} & \gamma_{i2} & \dots & \gamma_{ii} & \dots & \gamma_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \gamma_{n1} & \gamma_{n2} & \dots & \gamma_{ni} & \dots & \gamma_{nn} \end{pmatrix} \times \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_i & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & \lambda_n \end{pmatrix} \times \begin{pmatrix} \gamma_{11} & \gamma_{21} & \dots & \gamma_{i1} & \dots & \gamma_{n1} \\ \gamma_{12} & \gamma_{22} & \dots & \gamma_{i2} & \dots & \gamma_{n2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \gamma_{1i} & \gamma_{2i} & \dots & \gamma_{ii} & \dots & \gamma_{ni} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \gamma_{1n} & \gamma_{2n} & \dots & \gamma_{in} & \dots & \gamma_{nn} \end{pmatrix}$$

$$A = \Gamma \Lambda \Gamma'$$

Eq. 2.1

As matrizes Γ e Λ são obtidas como soluções das equações:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ii} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{ni} & \dots & a_{nn} \end{pmatrix} - \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_i & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & \lambda_n \end{pmatrix} = 0$$

$$|A - \Lambda| = 0$$

Eq. 2.2

e

$$\left[\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ii} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{ni} & \dots & a_{nn} \end{pmatrix} - \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_i & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & \lambda_n \end{pmatrix} \right] \times \begin{pmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1i} & \dots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \dots & \gamma_{2i} & \dots & \gamma_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \gamma_{i1} & \gamma_{i2} & \dots & \gamma_{ii} & \dots & \gamma_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \gamma_{n1} & \gamma_{n2} & \dots & \gamma_{ni} & \dots & \gamma_{nn} \end{pmatrix} = 0$$

$$(A - \Lambda)\Gamma = 0$$

Eq. 2.3

As matrizes decompostas nas análises de ordenação são simétricas e traduzem a associação (semelhança ou distância) entre objectos ou entre atributos, consoante o tipo de estudo em causa. As medidas de associação mais usadas são a soma dos produtos, a variância e a correlação, entre outros índices de distância entre pares de objectos ou entre pares de atributos. Numerosos autores desenvolveram índices próprios para traduzir essa associação e respectivas matrizes de associação; Legendre e Legendre (1984) e Podani (1990) recolhem os mais usados em ecologia. Legendre e Legendre (1984) organizam-nos numa chave dicotómica, de modo a facilitar a escolha do utilizador. Dever-se-á ter em conta a natureza do objecto determinante da sua semelhança e da estrutura ecológica a estudar, as restrições matemáticas devidas ao tipo de dados e aos cálculos a que a matriz é submetida na análise, e a disponibilidade do método em termos de programação.

Da análise espectral da matriz simétrica obtêm-se os valores $(\lambda_1, \lambda_2, \dots, \lambda_n)$ da matriz Λ , designados *valores próprios*, e os vectores $(\gamma_1, \gamma_2, \dots, \gamma_n)$, colunas da matriz Γ , designados *vectores próprios*. Assim definidos, os vectores próprios são ortogonais, isto é, não correlacionados entre si, e origem dum novo espaço multidimensional onde são reordenados os dados iniciais. Os *valores*

próprios representam a quantidade de variação recolhida por cada vector próprio da variação total da matriz de dados inicial. Como adiante se verá, tal tem implicações importantes no desenvolvimento das metodologias de análise de ordenação.

Neste resumo abordam-se apenas as metodologias mais comuns na criação e desenvolvimento de espaços multidimensionais onde se ordenam objectos e atributos investigados. As técnicas dividem-se em dois grupos, os quais se diferenciam pela avaliação dos atributos de forma directa ou indirecta na caracterização dos objectos investigados. Entre as metodologias de avaliação indirecta descrevem-se a análise de *componentes principais*, a análise *factorial*, a análise de *correspondências* e a análise de *coordenadas principais*. Entre as metodologias de avaliação directa de gradientes referimos a análise *discriminante* e a análise de *correlações canónicas*.

Prentice (1977) divide os métodos de ordenação em *métricos* e *não métricos*, apresentando os segundos como mais "potentes" em aplicações à ecologia. O autor descreve o caso de uma longa sucessão ecológica e a sua resposta a duas variáveis independentes: substrato pedológico e intensidade de pastoreio em pastos e comunidades vegetais mais complexas.

2.2.2.1 Análises indirectas de gradientes

As análises indirectas de gradientes aplicam-se quando não existe *à priori* nenhum conhecimento das formas e tendências da variação dos dados em investigação, fornecendo a primeira aproximação à estrutura e relações de objectos estudados, como base para aprofundamento de resultados por análise directa.

ANÁLISE DE COMPONENTES PRINCIPAIS

A análise de componentes principais foi desenvolvida por Hotelling (1933, citado por Frontier, 1982) e é uma das técnicas mais "potentes" de ordenação. Os *componentes principais* representam novas variáveis formadas a partir de combinações lineares das variáveis iniciais que, ao contrário destas, não são correlacionadas entre si. Como resultado, obtém-se a localização dos objectos em ordem a um espaço multidimensional, onde a variância se reparte por fracções independentes e identificadas com cada um dos eixos ou componentes. Os objectos, inicialmente num espaço *m*-dimensional, são transferidos para um novo espaço também *m*-dimensional, cujos eixos são completamente independentes entre si. Torna-se assim possível distinguir aspectos da estrutura dos dados que antes se encontravam encobertos pela dependência ou redundância entre os atributos iniciais. A análise de componentes principais é utilizada para atributos do tipo real, com dados quantitativos, e os seus resultados são válidos para distribuições do tipo multinormal.

Da matriz inicial ($X_{m \times n}$ - m atributos por n objectos) obtém-se outra ($S_{n \times n}$), quadrada e simétrica, que traduz a associação entre os n objectos em função da caracterização realizada pelos m atributos. Nos "packages" informáticos mais comuns, essa associação é expressa pela soma de produtos, pela covariância ou pela correlação (figura 2.5). A matriz de covariância é designada por matriz de *dados centrados* ($x_i - \bar{x}_i$), e a de correlação designada por matriz de *dados standartizados* ($\frac{(x_i - \bar{x}_i)}{s_x}$). Esta matriz é

decomposta, obtendo-se os respectivos valores e vectores próprios. Os primeiros indicam a variação que cada componente recolhe de entre a totalidade da variação total dos dados. Os segundos, quando ponderados com a raiz quadrada dos valores próprios, determinam os designados *componentes principais*, e dispõe-se em matriz diagonal. Estes componentes indicam a "composição" de cada eixo do novo espaço ortogonal pelas variáveis iniciais, e os seus valores são a correlação existente entre cada componente C_p e cada variável inicial x_j .

Da multiplicação da matriz de componentes principais pelos valores iniciais obtém-se a localização dos objectos no espaço agora definido. A interpretação dos componentes principais como combinação das variáveis iniciais nem sempre é fácil. A rotação dos eixos dos componentes principais por técnicas desenvolvidas por vários autores (Thurstone, 1954; Kaiser, 1958; Harman, 1967; Rummel, 1970; citado por Cíceri e col, 1977) permite colocar os eixos dos componentes principais numa nova posição, passível de ser mais facilmente identificável (Cíceri e col, 1977). Deste modo, os componentes inicialmente participados por todas as variáveis de uma forma equitativa, passam a ser definidos fortemente por um reduzido número delas e muito tenuamente pelas restantes. Para Gower (1967; citado por Prentice, 1977), a análise de componentes principais é um caso especial dos métodos de ordenação, onde são consideradas distâncias euclidianas como coeficientes de dissemelhança ou associação.

$$\begin{pmatrix} x_{ij} \end{pmatrix}^m \quad \begin{matrix} m - \text{número de variáveis} \\ n - \text{número de objectos} \\ x_{ij} - \text{valor da variável } j \text{ no objecto } i \end{matrix}$$

⇓ standartização (opcional)

$$\begin{pmatrix} r_{jj'} \end{pmatrix}^m \quad \begin{matrix} r_{jj'} = \frac{\text{COV}_{x_j x_{j'}}}{\sigma_{x_j} \sigma_{x_{j'}}} & \text{COV}_{x_j x_{j'}} - \text{covariância entre } j \text{ e } j' \\ & \sigma_{x_j}, \sigma_{x_{j'}} - \text{variância de } j \text{ e } j' \\ & r_{jj'} - \text{correlação entre } j \text{ e } j' \end{matrix}$$

⇓ análise espectral da matriz

$$\begin{pmatrix} l_{jp} \end{pmatrix}^m \quad \begin{matrix} x_j = \sum_{p=1}^m l_{jp} C_p & x_j - \text{valor da variável } j \\ & l_{jp} - \text{correlação entre } j \text{ e } p \\ & C_p - \text{componente } p \end{matrix}$$

⇓ rotação de eixos (opcional)

$$\begin{pmatrix} s_{ip} \end{pmatrix}^m \quad \begin{matrix} s_{ip} = \sum_{j=1}^m z_{ij} l_{jp} & z_{ij} - \text{valor standartizado da} \\ & & \text{variável } j \text{ no objecto } i \\ & s_{ip} - \text{valor do objecto } i \text{ para o} \\ & & \text{componente } p \end{matrix}$$

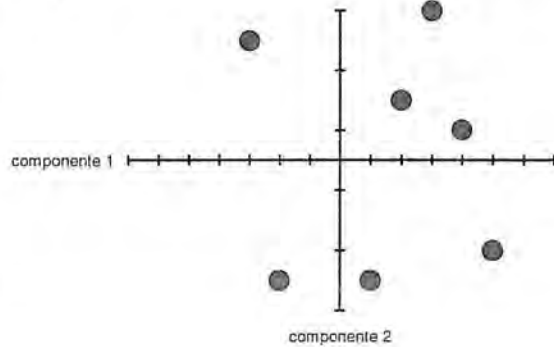


Figura 2.5 - Desenvolvimento-base da análise de componentes principais (modelo com utilização da matriz de correlação).

ANÁLISE FACTORIAL

A análise factorial corresponde a uma análise de componentes principais na qual se pretende operar apenas sobre a variação dos dados que é comum a variáveis consideradas representativas. Na análise precedente, a totalidade desta variação é transferida para os componentes principais. Na análise factorial analisa-se e representa-se apenas a variação que por um qualquer critério se define como comum às variáveis de base. Embora seja considerada um caso particular da análise de componentes principais, a sua base teórica teve origem em estudos de ciências sociais, no princípio do século (Spearman, 1904, citado por Lefebvre, 1983).

Em função do conhecimento de causa e do seu significado, o investigador decide os "factores" a incluir na análise, rejeitando os que não apresentam uma significação lógica.

Utilizando um critério que permite a selecção da variação comum às variáveis apresentadas, calcula-se a respectiva matriz simétrica. No caso de se optar pela matriz de correlação, os valores da sua diagonal poderão ser inferiores ou iguais a 1, ao contrário do que se passa no caso anterior. A correlação entre os factores e as variáveis iniciais calcula-se tendo em conta uma quantidade adicional de variação exclusiva dessa mesma variável inicial. A matriz final de localização dos objectos no novo espaço k -dimensional criado, é naturalmente mais reduzida do que aquela concebida na análise de componentes principais. Esta redução da quantidade de informação em relação à inicial, sem perda de valor explicativo, é muitas vezes o objectivo da utilização da análise factorial, onde a relação de ortogonalidade entre os eixos do novo espaço constituído permite a melhor interpretação da estrutura dos dados em estudo.

$$\begin{array}{l}
 \left(\begin{array}{c} x_{ij} \\ \vdots \\ x_{ij} \end{array} \right)_n^m \quad \begin{array}{l} m - \text{número de variáveis} \\ n - \text{número de objectos} \\ x_{ij} - \text{valor da variável } j \text{ no objecto } i \end{array} \\
 \Downarrow \text{standartização (opcional)} \\
 \left(\begin{array}{c} r_{jj'} \\ \vdots \\ r_{jj'} \end{array} \right)_m^m \quad \begin{array}{l} r_{jj'} = \frac{\text{cov}_{x_j x_{j'}}}{\sigma_{x_j} \sigma_{x_{j'}}} \quad \begin{array}{l} \text{cov}_{x_j x_{j'}} - \text{covariância entre } j \text{ e } j' \\ \sigma_{x_j}, \sigma_{x_{j'}} - \text{variância de } j \text{ e } j' \end{array} \\ r_{jj'} - \text{correlação entre } j \text{ e } j' \end{array} \\
 \Downarrow \text{análise espectral da matriz} \\
 \left(\begin{array}{c} l_{jp} \\ \vdots \\ l_{jp} \end{array} \right)_m^k \quad \begin{array}{l} x_j = \sum_{p=1}^k l_{jp} f_p + \varepsilon_j \quad \begin{array}{l} x_j - \text{valor da variável } j \\ l_{jp} - \text{correlação entre } j \text{ e } p \\ f_p - \text{factor } p \\ \varepsilon_j - \text{variação única de } j \end{array} \\ \Downarrow \text{rotação de eixos (opcional)} \\
 \left(\begin{array}{c} s_{ip} \\ \vdots \\ s_{ip} \end{array} \right)_n^k \quad \begin{array}{l} s_{ip} = \frac{\sum_{j=1}^k z_{ij} l_{jp}}{\lambda_p} \quad \begin{array}{l} z_{ij} - \text{valor standartizado da} \\ \text{variável } j \text{ do objecto } i \\ s_{ip} - \text{valor do objecto } i \\ \text{para o componente } p \\ \lambda_p - \text{valor próprio } p \end{array} \end{array} \\
 \Downarrow
 \end{array}$$

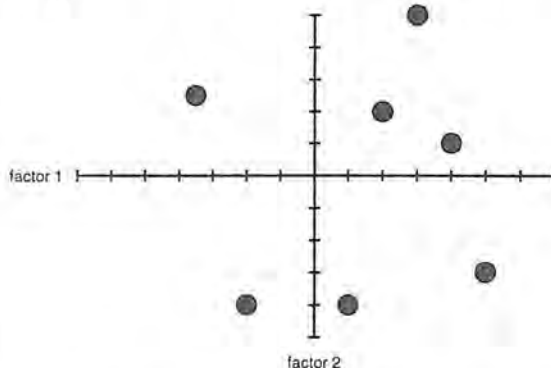


Figura 2.6 - Desenvolvimento-base da análise factorial.

ANÁLISE DE CORRESPONDÊNCIAS

A análise de correspondências aparece para responder a duas características das estruturas dos sistemas naturais não tidas em conta no tratamento numérico das análises anteriormente descritas: atributos do tipo qualitativo (nominais ou binários, aos quais corresponde uma matriz de frequências) e/ou com distribuições de frequência unimodais. Neste tipo de análise, a ordenação de objectos e de atributos é feita simultaneamente, o que facilita a clarificação das suas relações. A sua origem remonta aos anos 30, e desde então é descrita por numerosos autores sob as mais diversas denominações. Os autores franceses referem-na como análise factorial de correspondências e os ingleses por análise de tabelas de contingência, ponderação recíproca, técnica R-Q, ordenação recíproca e análise de homogeneidade (Legendre e Legendre, 1984).

Os dados de uma matriz inicial de frequências são ponderados de modo a que, quando com valores muito diferentes, não comprometam a sua correcta interpretação. A nova matriz corrigida, dita centrada, é multiplicada pela sua transversa, obtendo-se uma matriz quadrada simétrica, indicadora da associação dos objectos ou dos atributos, consoante o sentido dessa multiplicação (a multiplicação de matrizes não é comutativa). De essa matriz são determinados os respectivos valores e vectores próprios. A localização dos atributos e dos objectos em ordem ao espaço formado pelos k -novos factores é obtida pelos vectores próprios e são passíveis de ser representados num mesmo espaço k -

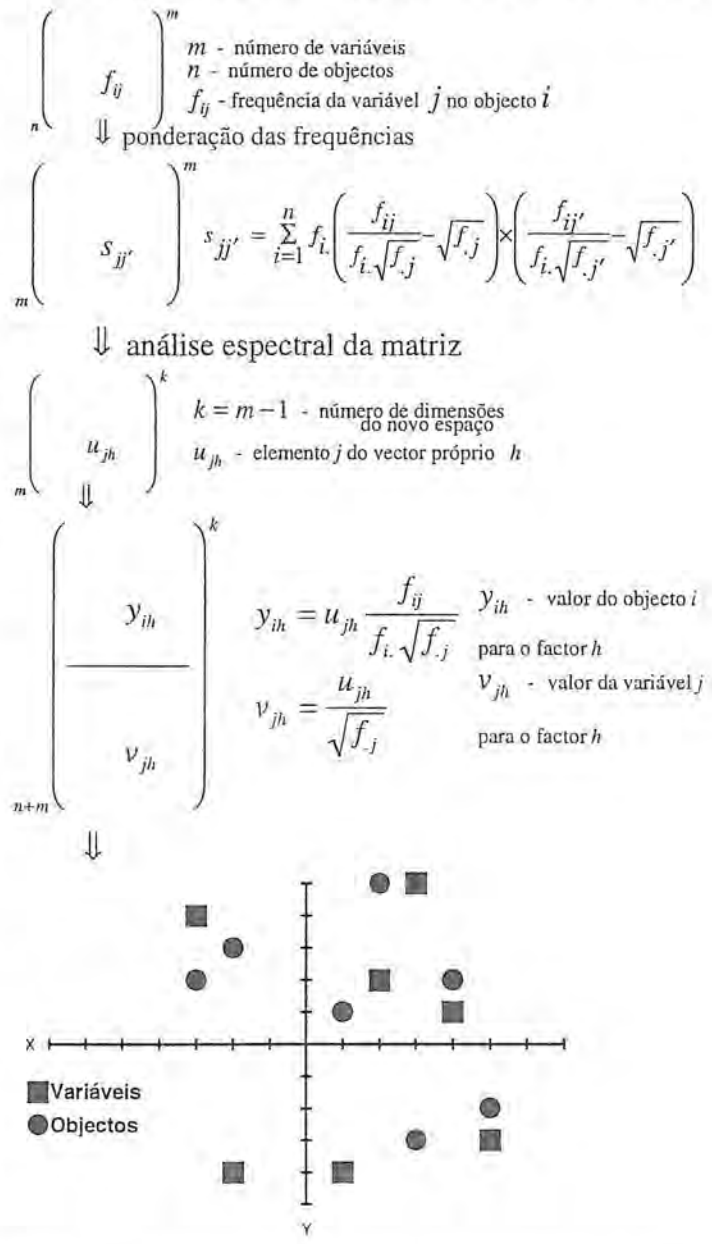


Figura 2.7 - Desenvolvimento-base da análise de correspondências.

dimensional. Quando a matriz simétrica de base corresponde à associação dos atributos, considera-se que a representação é feita no *baricentro* dos objectos, e no caso dessa matriz representar a associação dos objectos, consideram-se representados no *baricentro* dos atributos.

Uma análise corrigida de correspondências ("detrended correspondence analysis") é descrita como uma variação da análise descrita antes, cujo objectivo é desfazer algum efeito de concentração dos objectos ou dos atributos nos extremos dos eixos de ordenação, bem como a atenuação do efeito de atributos raros em objectos com poucos atributos (Fry, 1993).

ANÁLISE DE COORDENADAS PRINCIPAIS

Desenvolvida por Gower (1966; citado por Podani, 1990), e ao contrário das anteriores, a análise de coordenadas principais não utiliza a soma dos produtos, as covariâncias ou as correlações para avaliar a associação entre objectos ou variáveis, mas índices específicos, ditos de distância ou de semelhança. Hoje em dia, os "packages" mais vulgares disponibilizam ao utilizador várias dezenas de índices para eleger, de acordo com o tipo e estrutura dos dados em causa. Deste modo, e ao contrário das análises anteriores, é permitida a utilização de vários tipos de dados desde que se escolha um índice adequado.

Da matriz inicial de objectos por atributos, e através de um adequado índice de associação, é calculada uma matriz simétrica de "distâncias" entre os objectos. Os seus valores são transformados pela fórmula apresentada na figura 2.8, a qual, preservando as relações de semelhança entre os objectos, permite situar a origem do novo sistema de eixos no centróide dos atributos estudados. A matriz resultante é decomposta em valores e vectores próprios, sendo o número destes, e consequentemente o dos factores (a dimensão do novo espaço

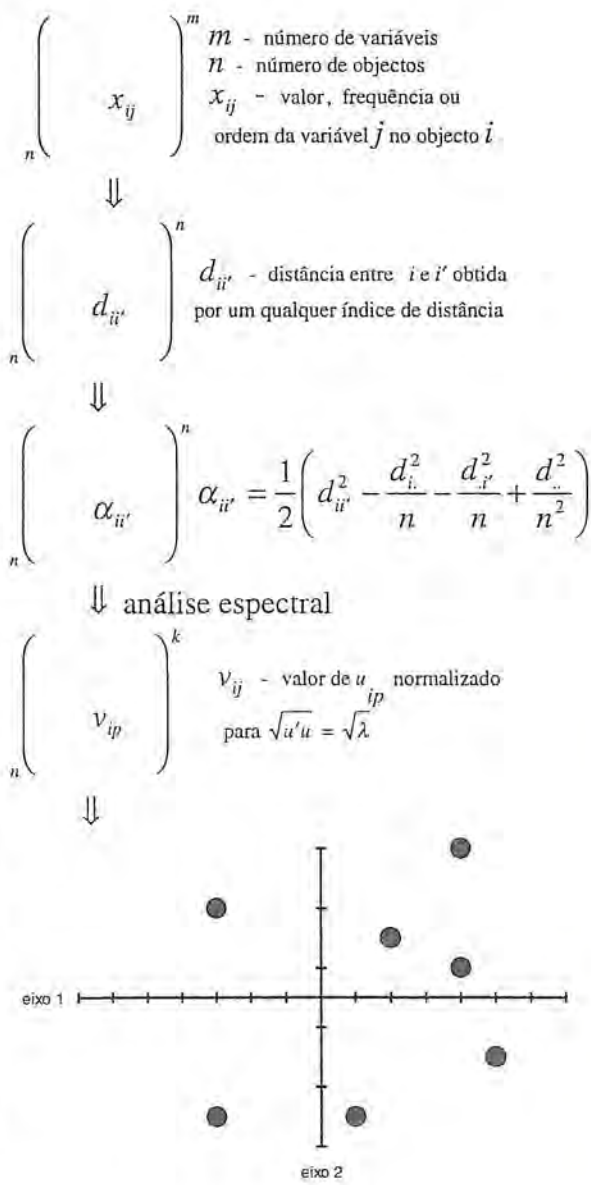


Figura 2.8 - Desenvolvimento-base da análise de coordenadas principais.

ortogonal), dependente da relação entre o número de objectos e o número de atributos. Quando o número de atributos é superior e a matriz não degenerada, o número máximo de valores próprios é igual ao de objectos. Quando o número de objectos é superior, o de valores próprios é inferior ao de atributos, como consequência dos valores da matriz serem centrados. A matriz de localização dos objectos obtêm-se pela normalização dos vectores próprios em relação à raiz quadrada dos valores próprios, resultando daí novas coordenadas para os objectos, ordenados num espaço k -dimensional.

O êxito da análise efectuada depende duma correcta escolha da medida de associação. Ao contrário da análise de componentes principais, a comparticipação de cada variável para a composição dos eixos do novo espaço não é determinada directamente mas por análise posterior. Leech e col (1991) utilizaram a análise de coordenadas principais para conhecer a relação entre várias espécies arbóreas, em função das respectivas equações de volume, usando o T^2 de Hotelling como coeficiente de associação. O método revelou-se promissor, recomendando os seus autores a utilização desta técnica como complementar de outras.

2.2.2.2 Análise directa de gradientes

ANÁLISE DAS CORRELAÇÕES CANÓNICAS

Na análise multidimensional de dados ecológicos, a identificação da "responsabilidade" de factores ambientais no desenvolvimento de qualquer fenómeno é feita com base nas *correlações canónicas*. Esses parâmetros permitem estimar a quantidade e a qualidade das ligações entre dois grupos de variáveis (ambientais e bióticas, dependentes e independentes, etc), da mesma forma que o faz a correlação simples entre duas variáveis e a correlação múltipla entre uma variável e um grupo de variáveis/atributos. O método das correlações canónicas é a generalização da regressão múltipla desenvolvida por Hotelling (1936; citado por Legendre e Legendre, 1984),

sendo esta e a regressão linear, consideradas casos particulares das correlações canónicas.

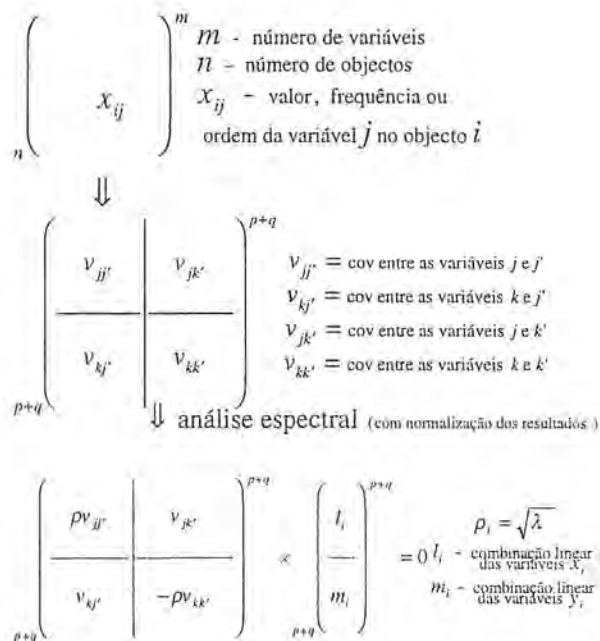


Figura 2.9 - Desenvolvimento-base da análise das correlações canónicas.

Para uma matriz inicial de $p + q$ variáveis por n objectos calcula-se a respectiva matriz da soma dos produtos, das covariâncias ou das correlações. Essa matriz é composta por quatro blocos:

duas correspondem às relações entre as variáveis x e entre as variáveis y , e as duas restantes, inversa uma da outra, correspondem às "relações" entre variáveis x e y . A correlação canónica entre os dois grupos de variáveis é a máxima correlação existente entre as combinações lineares intragrupo dessas variáveis (Lefebvre, 1983). O seu cálculo passa pela maximização da variação entre os grupos de variáveis, minimizando-a dentro desses grupos. A solução consegue-se pela análise espectral da matriz daí resultante, da qual derivam p soluções possíveis. As correlações canónicas (ρ_i) são a raiz quadrada dos valores próprios obtidos pela análise espectral. Dado a multiplicação de matrizes não ser comutativa, a cada um destes valores próprios diferentes correspondem dois vectores próprios, combinações lineares das variáveis de origem x e y . Essas combinações lineares determinam então duas novas variáveis ditas canónicas, l e m , substituindo as $p + q$ variáveis iniciais (Lefebvre, 1983).

$$l = \mathbf{l}'\mathbf{x} = (l_1x_1 + l_2x_2 + \dots + l_px_p) \text{ e } m = \mathbf{m}'\mathbf{y} = (m_1y_1 + m_2y_2 + \dots + m_qy_q) \quad \text{Eq. 2.4}$$

Cada um desses p conjuntos de dois vectores próprios são normalizados, tornando igual à unidade a variação das variáveis canónicas. O resultado corresponde às coordenadas dos objectos em cada um dos dois sistemas de eixos principais.

Teoricamente, as correlações canónicas são um instrumento na escolha duma ordenação das matrizes duplas de dados (dados de vegetação e dados físicos, por exemplo). No entanto, Gauch (1982) demonstra que, tal como nos componentes principais, apenas se aplica a variáveis de resposta linear, o que nem sempre acontece em ecologia. Por outro lado, ao contrário dos coeficientes de determinação simples e múltiplos, o quadrado das correlações canónicas não traduz a proporção da variação que uma variável canónica explica da sua complementar; para tal foram desenvolvidos índices de redundância por Stewart e Love (1968) e Cooley e Lohnes (1971; citado por Legendre e Legendre, 1984). Ao contrário do coeficiente de determinação simples, este índice não é necessariamente simétrico, pelo facto das combinações das variáveis no primeiro não corresponderem ao segundo conjunto. Feoli e Orloci (1979) propõe uma análise de concentração aplicada a tabelas de objectos por atributos divididos em blocos. Baseada nas correlações canónicas, esta estrutura em bloco é testada para avaliação da significância das diferenças nos blocos constituídos. Uma análise vectorial às frequências de cada par objecto-atributo permite estimar correlações canónicas e coordenadas canónicas para cada grupo de objectos e atributos. Os quadrantes obtidos podem então ser confrontados com variáveis exteriores, passíveis de identificar os maiores traços da variação dos dados.

ANÁLISE DISCRIMINANTE

A análise discriminante pressupõe um pré-conhecimento dos dados que permite, não uma divisão dos atributos como nas correlações canónicas, mas sim dos objectos, em duas ou mais classes. Tal divisão baseia-se numa prévia análise de classificação ou em caracteres intrínsecos dos dados. Em ecologia, a proximidade no terreno ou agrupamento dos objectos servem a essa pré-classificação, pretendendo-se então distinguir os atributos que melhor acompanham a variação dos objectos/indivíduos ao longo dos gradientes dum território. As primeiras aplicações à ecologia são simultâneas ao desenvolvimento das funções discriminantes de Fisher (1936) e do coeficiente de distância multivariada de Mahalanobis (1936; citados por Ludwig e Reynolds, 1988).

A análise discriminante de uma matriz de dupla entrada, cujos elementos x_{ijk} se referem ao valor do objecto i , pertencente ao grupo k e caracterizado pela variável j , desenvolve-se pela maximização das diferenças entre grupos. Tal procedimento implica a decomposição da matriz que descreve a relação entre a dispersão dos valores intergrupo e dos intragrupo. Os vectores próprios obtidos, ao contrário das outras análises, não são ortogonais, pela razão da sua matriz de origem não ser simétrica. No entanto, providenciam um novo referencial aos objectos, os quais compõem elipsóides correspondentes aos grupos dos objectos originais. Tais formas elipsóides transformam-se em circulares pela normalização dos vectores próprios (figura 2.10).

$$\begin{aligned}
 & \left(\begin{array}{c} x_{ijk} \\ \vdots \\ x_{ijk} \end{array} \right)^m \quad \begin{array}{l} m - n^\circ \text{ de variáveis} \\ n - n^\circ \text{ de objectos, organizados em } O \text{ grupos} \\ x_{ijk} - \text{valor da variável } j \text{ no objecto } i, \text{ do grupo } O \end{array} \\
 & \Downarrow \\
 & \left(\begin{array}{c} t_{jj'} \\ \vdots \\ t_{jj'} \end{array} \right)^m \quad t_{jj'} = \sum_{k=1}^o \sum_{i=1}^n (x_{ijk} - \bar{x}_j)(x_{ij'k} - \bar{x}_{j'}) \\
 & \text{e} \\
 & \left(\begin{array}{c} w_{jj'k} \\ \vdots \\ w_{jj'k} \end{array} \right)^m \quad w_{jj'k} = \sum_{i=1}^n (x_{ijk} - \bar{x}_{jk})(x_{ij'k} - \bar{x}_{j'k}) \\
 & \Downarrow \\
 & \left(\begin{array}{c} b_{jj'} \\ \vdots \\ b_{jj'} \end{array} \right)^m \quad b_{jj'} = t_{jj'} - \sum_{k=1}^o w_{jj'k} \\
 & \Downarrow \text{ análise espectral (com normalização dos resultados)} \\
 & \left(\begin{array}{c} c_{ijk} \\ \vdots \\ c_{ijk} \end{array} \right)^p \quad \begin{array}{l} c_{ij} = u_{ip} \text{ normalizado} \\ \text{para } \sqrt{u'u} = \sqrt{\lambda} \end{array} \\
 & \Downarrow
 \end{aligned}$$

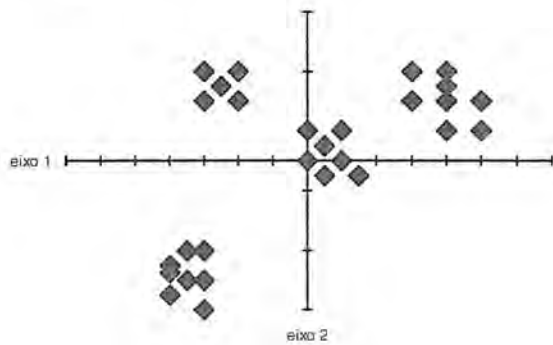


Figura 2.10 - Desenvolvimento-base da análise discriminante.

A análise discriminante proporciona a obtenção de eixos sobre os quais a dispersão dos centros dos grupos de objectos é a maior possível. Em resumo, o seu procedimento é em tudo igual ao dos componentes principais, nos quais se procura essa máxima dispersão para a totalidade dos objectos. Szaro e King (1990) utilizaram a análise discriminante para estudar a variabilidade das características físicas do meio, entre e dentro de comunidades ripícolas dos Estados Unidos. Tais comunidades foram estabelecidas pela análise da vegetação de 153 locais que permitiram a definição de 28 comunidades por uma análise hierárquica aglomerativa. Tal agrupamento permitiu o estabelecimento de variáveis canónicas, como combinações lineares das variáveis iniciais, não correlacionadas entre si, e com a maior correlação possível com as comunidades. Uma análise discriminante do tipo "stepwise" permitiu interpretar os gradientes das comunidades estabelecidas. Alvarez e col (1989) procederam exactamente ao contrário; o território de um vale da Cordilheira Cantábrica foi classificado em função de suas variáveis geofísicas e da sua relação com a utilização dominante. Tal agrupamento foi significativamente discriminado, utilizando como variáveis as espécies observadas em inventários florísticos realizados ao longo do território de estudo. Vos e Stortelder (1992) discriminam classificações hierárquicas de inventários florísticos e perfis de humus em relação a variáveis ambientais, constatando a importância de factores mais gerais nos níveis mais altos de classificação, e o aparecimento de factores mais detalhados como explicativos das divisões de níveis hierárquicos mais baixos.

Vários programas de computador estão hoje disponíveis para a realização das análises de ordenação referidas acima. De entre eles, alguns citados ao longo do texto, destacam-se *CANOCO* (Hill, 1979), *DISCRIM* (Ter Braak, 1982) e *SYNTAX* (Podani, 1990).

CANOCO é um desenvolvimento de outro programa, *DECORAMA* (Hill, 1979); está preparado para calcular Componentes Principais, Coordenadas Principais, Correspondências, Correlações Canónicas e respectivas variações.

DISCRIM é uma modificação de outro programa, *TWINSPAN* (Hill, 1979), já referido como método de classificação; está preparado para discriminar classificações hierárquicas, classificando atributos em função da sua relação com o desenvolvimento dessas classificações.

SYNTAX (Podani, 1979), de entre as análises de ordenação, calcula Componentes Principais, Coordenadas Principais e Correspondências; este programa destaca-se dos demais pela quantidade de índices de semelhança que disponibiliza na análise de coordenadas principais, bem como pela possibilidade de avaliar semelhanças entre ordenações realizadas por diferentes métodos. A comparação de ordenações pode assumir a forma dum valor que expressa a sua semelhança, ou a construção de uma nova ordenação de consenso entre as diversas em confronto.

2.3 ENTROPIA E INFORMAÇÃO: significado e utilização

Entropia, termo derivado do grego "entrepein" que significa *transformar em*, relaciona-se com uma das leis fundamentais da física, a 2ª lei da termodinâmica de Boltzman, associando-se a desordem. Essa lei estabelece que todos os processos espontâneos de transformação de energia em sistemas fechados levam a formas de energia mais dispersas e por isso mais degradadas e menos disponíveis. Um estado final de completa homogeneidade - estado de equilíbrio estático - atinge-se quando todas as diferenças de gradientes energéticos são igualados. Considerando que a entropia é baixa nos sistemas ordenados e alta nos sistemas desordenados, o incremento da entropia está associada à estabilização, o estado " Ω " para o qual tendem os sistemas naturais à medida que evoluem. Entropia é, portanto, uma medida do estado dum sistema, da sua homogeneidade e irreversibilidade, um processo que toma apenas uma direcção e cujo incremento está associado à imprevisibilidade. Tal conduz ao paradoxo de que a imprevisibilidade não será uma característica da instabilidade dos sistemas naturais.

A sua quantificação matemática foi feita por Shannon, dos laboratórios de Bell Telephone, em 1948, e hoje é uma equação bem conhecida (Legendre e Legendre, 1984), avaliando-se pelo logaritmo das proporções entre os diversos estádios adquiridos por um sistema e a totalidade dos estádios possíveis nesse mesmo sistema. A equação fundamental de Shannon descreve simbolicamente uma quantidade que se desloca de um emissor para um receptor.

$$H = -K \sum_{i=1}^n p_i \log p_i \quad \text{Eq. 2.5}$$

Esta função é explicada por Daget (1979) como definição do conceito de *indeterminação*. Segundo este autor, *indeterminação* é o logaritmo de base 2 do número de casos possíveis com igual probabilidade de ocorrência. Como exemplo demonstrativo, refere a indeterminação associada ao resultado de diversas capturas realizadas entre uma população de N indivíduos. Para tal, a função terá de cumprir os requisitos:

$$\text{Indeterminação} = f(N) \cdot \begin{cases} f(1) = 0 \\ f(N^2) = 2f(N) \end{cases} \Rightarrow f(N) = \log_2 N$$

Essa função é a função logarítmica. A base 2 deriva de se pretender exprimir o valor da função como a probabilidade entre dois resultados (alternativas simples, presença/ausência, etc.). De outra forma, a utilização do logaritmo decimal por exemplo, implica considerar 10 possibilidades diferentes de resultado. O uso do logaritmo de base 2 conduz ao uso da unidade "bit", abreviação inglesa de unidade binária ("*binary digit*"). No entanto, o valor da indeterminação é também

expresso como a soma do contributo de cada indivíduo para o valor total dessa população. O seu valor é expresso pela equação 2.6.

$$\text{Indeterminação} = - \sum_{i=1}^N \frac{1}{N} \log_2 \frac{1}{N} \quad \text{Eq. 2.6}$$

Como dificilmente a probabilidade é igual para todos os constituintes de uma população e/ou essa população nem sempre é dividida em espécies ou outro qualquer qualificativo, chegar-se-ia então à função proposta por Shannon (equação 2.5).

Este conceito apresenta uma marcada analogia com o de *informação*. Informação, representa a diferença existente entre a *indeterminação* associada a uma mensagem e a sua anulação após o seu conhecimento, cujo valor é igual a zero ($1 \times \log 1 = 0$) e que portanto deve-se considerar igual à quantidade de indeterminação.

$$I = \log_2 \frac{P}{P_0} \quad \text{Eq. 2.7}$$

Tal como o conceito de entropia física, o conceito de informação quantifica a distância entre um sistema perfeitamente previsível (entropia mínima - probabilidade de ocorrência igual a 1 - sistemas determinísticos) e um sistema qualquer cujos resultados mais baixos se atingem quando os objectos se repartem por várias classes de igual probabilidade (ou frequência relativa). A informação, originalmente teorizada por Shannon e Weaver (1949; citado por Daget, 1979), reflecte a incerteza de um acontecimento e a sua redução por uma mensagem. A quantidade de informação de uma mensagem é expressa matematicamente pelo número de escolhas entre duas possibilidades de probabilidade igual, ou decisões binárias ("bits").

A aplicação da teoria da informação à biologia influenciou a sua base conceptual e, em alguns casos, a sua metodologia (Orlóci, 1991a). As bases teóricas do conceito

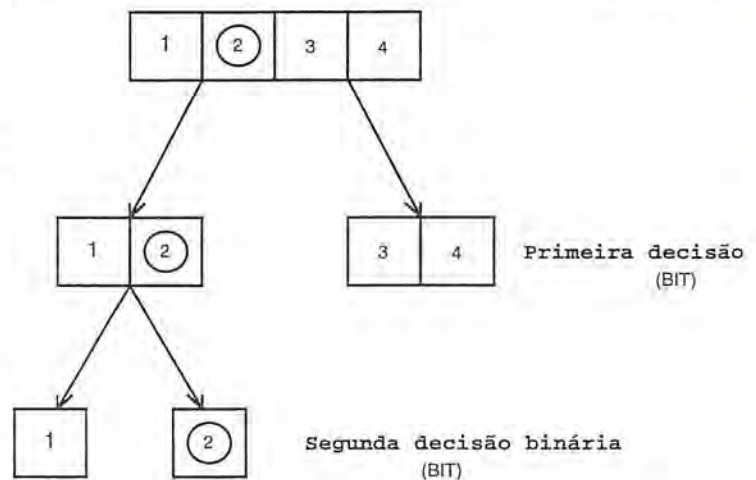


Figura 2.11 - Decomposição da informação em decisões binárias entre 4 símbolos (N) (adapt. Naveh e Lieberman, 1983)

de informação são descritas como úteis na quantificação das propriedades estruturais, diversidade, mutualidade e equivocação. A teoria da informação oferece a vantagem da sua identificação universal na descrição das estruturas e relações entre populações e comunidades. As metodologias

aplicam-se à modelação de sistemas, estimação de diversidade e análise estatística de complexos de dados ecológicos.

Em Frontier (1982), o conceito de entropia (H) é apresentado como parâmetro de dispersão da estatística descritiva das distribuições de frequência de atributos do tipo cardinal. O mesmo autor afirma que para dados qualitativos, o equivalente à variância de um atributo quantitativo i , é a quantidade de entropia $H(i)$ calculada pela equação de Shannon (equação 2.5), com as proporções dos valores não-ordenados de uma variável. A inexistência nestes casos, de qualquer relação de ordem entre os valores, leva a que se tome em conta apenas as frequências de cada classe.

Da mesma forma que a entropia ou a informação traduz a estrutura de base de um conjunto de objectos ao avaliar a sua heterogeneidade, serve também para inferir sobre a variação conjunta de atributos de distribuições do tipo cardinal (qualitativa). A figura 2.12 pretende ilustrar isso mesmo. A informação partilhada por duas variáveis calcula-se tendo em conta a sua informação total $H(i,k)$ e a informação de cada uma individualmente:

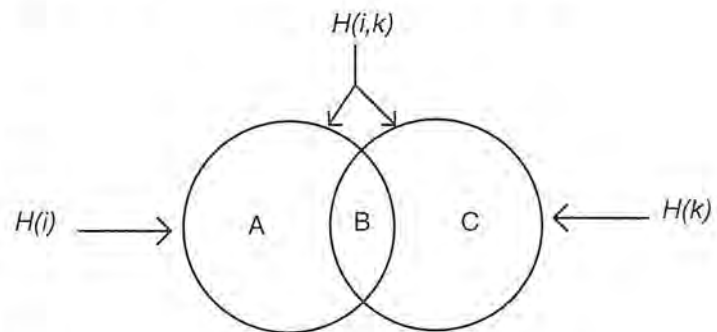


Figura 2.12 - Representação gráfica da relação entre os conceitos de entropia e informação mútua (adapt. Daget e Godron, 1982)

$$B = H(i) + H(k) - H(i,k) \quad \text{Eq. 2.8}$$

Pela diferença obtém-se A e C, a entropia que é realmente partilhada pelas duas variáveis ou classificações.

$$A = H(i|k) = H(i,k) - H(k) \quad \text{Eq. 2.9}$$

$$C = H(k|i) = H(i,k) - H(i) \quad \text{Eq. 2.10}$$

À quantidade de variação partilhada, Daget e Godron (1982) referem-se como sendo a informação mútua entre A e C. Em fitossociologia, ela exprime a sociabilidade da espécie A com a espécie C. As espécies, como variáveis do tipo cardinal, têm na entropia um índice de diversidade por excelência. Margalef (1957; citado por Daget, 1979), afirma ser essa diversidade, a informação básica transmitida pelo meio sobre a composição e estrutura de um ecossistema.

A análise das relações ao nível da comunidade e população é um dos campos mais desenvolvidos das aplicações da teoria de informação à biologia. Margalef (1957; citado por Daget,

1979) demonstra como este valor pôde ser descrito em ecologia; uma unidade de amostragem que compreenda Q indivíduos divididos e classificados em N espécies, com os quantitativos de $q_1, q_2, q_3, \dots, q_n$, pode ser entendida como uma mensagem que fornece informação sobre a composição e estrutura da população de onde ela é destacada. A natureza deste conceito e o seu cálculo, faz com que seja considerado um valioso índice de diversidade, tomando o nome de índice de diversidade de Shannon. Em ecologia, é utilizado para hierarquizar variáveis em função da informação comum à variabilidade total dos dados.

Orlóci (1976) propõe a substituição da tradicional soma de quadrados por critérios de informação, mais adequados a dados qualitativos relativos a vegetação, quando se pretende avaliar a importância de um atributo na informação mútua de um conjunto de variáveis. O autor propõe a designação de informação mútua específica como a diferença entre a

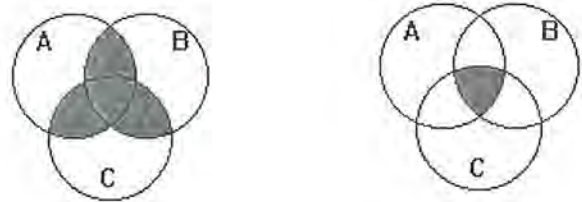


Figura 2.13 - Informação interactiva (área sombreada à esquerda) e informação mútua (área sombreada à direita) de uma distribuição de frequência com três dimensões. (ORLÓCI, 1991a)

informação do conjunto de variáveis e a informação desse conjunto sem a participação dessa variável. Lepart e Debussche (1980) também propõem critérios de informação para a avaliação da eficácia das variáveis ambientais na distribuição de comunidades de *Lanuginosa*, bem como para a definição do perfil de eficiência de cada variável ao longo dos diversos estádios que atravessa, no qual diferencia a eficácia desse estádio e a contribuição para essa eficiência.

De Pablo e col (1987) utilizam parâmetros de informação para avaliar três tipos diferentes de classificação dum território, em função de atributos climáticos, litológicos, topográficos e de vegetação e uso do solo. A eficácia de uma cartografia devida a determinada variável é expressa pela diferença entre entropia da sectorização do território e a entropia dessa sectorização atendendo à distribuição dessa variável. O quociente entre essa eficácia e a entropia da sectorização, tendo em conta a distribuição da variável, indica a eficácia com que essa variável é utilizada para a sectorização do território. O valor médio da eficácia das variáveis que são utilizadas no reconhecimento de um território, permitirá avaliar a cartografia que melhor traduz a estrutura dos atributos caracterizadores de um território. De Pablo e col (1988) destacam a relação entre atributo indicador e objecto indicado, servindo-se de parâmetros da teoria da informação. Para os autores, a informação mútua representa a informação sem distorção transmitida pelos estádios das variáveis que caracterizam um território, sobre os sectores que constituem esse território. No seu estudo é

proposto a redefinição dos estádios das variáveis e/ou sectores de um território, como forma de minimizar essa distorção, maximizando a informação transmitida.

Gómez Sal e Muñoz-Yanguas (1990) utilizam parâmetros de informação para analisar a coincidência de elementos paisagísticos, definindo modelos do tipo tecelar do mosaico territorial de uma região da Cordilheira Cantábrica. Gómez Sal e col (1992) utilizam a informação mútua para avaliar a importância de variáveis físicas, de usos e índices faunísticos na avaliação do Valor Natural e Complexidade Estrutural de uma paisagem. Gómez Sal e col (1993) utilizam a informação mútua para avaliar a semelhança entre a distribuição espacial de usos de um território em três datas distintas.

CAPÍTULO 3

ECOLOGIA DE TRÁS-OS-MONTES E ALTO DOURO

3.1 APRESENTAÇÃO DO TERRITÓRIO

Trás-os-Montes e Alto Douro é a província que ocupa o Nordeste do território de Portugal, com aproximadamente 21600 km², entre 41° e 42° de latitude Norte e 8° e 6° de longitude Oeste. É um espaço na fronteira entre as biogeografias euro-siberianas e mediterrâneas, onde a orografia assume particular importância na diversificação das suas características ambientais. Drenado pela bacia hidrográfica do rio Douro, o território é marcado pelo vale cavado desse rio, particularmente na sua região Este e Sul, com altitudes entre 50 e 500 metros e condições tipicamente mediterrâneas. Vários dos afluentes deste rio cruzam a província, com uma direcção dominante Norte/Sul, sendo os rios Tâmega, Corgo, Tua e Sabor, os mais importantes. Entre estes rios, desenvolvem-se pequenas áreas planálticas de base xistosa e altitude média entre 600 e 800 metros, entrecortadas por numerosos sistemas montanhosos, na sua maioria de origem granítica. Mais ou menos importantes, estas elevações apresentam características típicas dos complexos serranos galaico-durienses, ou seja, altitudes situadas entre os 1000 e 1500 metros e clima temperado húmido; entre elas, é possível distinguir as de maior influência atlântica (Barroso, Larouco, Cabreira e Montemuro), as de carácter intermédio (Leomil, Alvão, Marão e Padrela), e as de maior influência continental (Bornes, Nogueira, Coroa, Montesinho e Mogadouro).

O Norte Transmontano é uma das "Divisões Fundamentais da Terra Portuguesa" definidas pelo geógrafo Orlando Ribeiro (Ribeiro, 1987), juntamente com o Norte Atlântico e o Sul. Esta divisão - Norte Transmontano - prolonga-se ao Sul da província até ao sopé da Cordilheira Central, e separa-se do Norte Atlântico pelas montanhas que atenuam as influências marítimas (Montemuro, Marão, Alvão, Padrela e Barroso). Ribeiro (1987) caracteriza esta região como "altas plataformas onduladas, cortadas por vales e bacias muito profundas". A paisagem possui grandes horizontes, onde vastos campos de cereais se apresentam pontilhados por afloramentos rochosos e bosquetes de carvalho negral (*Quercus pirenaica*), ulmeiro (*Ulmus nigra*) e castanheiro (*Castanea sativa*). Na utilização agrária do território predomina a rotação centeio/pousio, à qual se junta a batata nas regiões mais favoráveis, numa paisagem que lembra muitas vezes as regiões agrícolas da Europa Central. As depressões mais húmidas são aproveitadas como pastagem para ruminantes (lameiros). O pinhal e as culturas regadas acompanham a influência atlântica na região.

A partir do século XVII, o Homem começa a colonização das encostas declivosas do Douro e seus afluentes, construindo socalcos nos quais plantou a vinha que está na origem do Vinho do Porto. O aparecimento da filoxera, praga que nos finais do século XVIII dizimou a quase totalidade das áreas vinícolas, alterou este quadro; a recolonização com videiras resistentes deixou as zonas baixas para a oliveira e amendoeira entretanto introduzidas, e os cerros mais inóspitos com matos ou bosquetes de resinosas (mortórios) (Ribeiro, 1987). No entanto, esta paisagem não deixa de constituir uma das mais impressionantes "infra-estruturas agrárias" de Portugal.

A utilização da província de Trás-os-Montes e Alto Douro como objecto de investigação permite o estudo de um território que concentra uma alta diversidade interna, fruto sobretudo do cruzamento de influências atlânticas, mediterrâneas e continentais, e de litologias graníticas e xistosas. Como resultado, a sua paisagem rural apresenta-se como um mosaico de alto valor informativo e de conservação, muito susceptível às transformações da Política Agrícola Comum. A dimensão da área em estudo reflete por um lado, um valor mínimo que dê sentido aos factores ecológicos em estudo, e por outro, um valor máximo que tenha em atenção a homogeneidade dos dados disponíveis para tal tarefa. A publicação da 2ª versão preliminar da Carta dos Solos, Carta de Utilização Actual do Solo e Carta de Uso Potencial do Solo do Nordeste de Portugal (Agroconsultores e COBA, 1990) foi fundamental na disponibilização de informação relativa à ecologia do território, tanto dos seus principais factores, como da sua repartição espacial. Tal obra abrange o território da Província, a qual se estende pelos distritos de Vila Real e Bragança, e ainda os concelhos limítrofes dos distritos de Viseu e Guarda.

A recolha dos dados deste estudo foi realizada sobre cartografia temática existente e disponível, de uma forma sistemática. Os locais de amostragem localizam-se ao longo de oito eixos de longitude da Rede Geodésica Internacional (meridiano de Greenwich) com 15' de intervalo entre si, que percorrem a área de estudo - província de Trás-os-Montes e Alto Douro; dentro dos eixos de longitude foram inventariados pontos com intervalos de um minuto de latitude entre si.

A utilização da Rede Geodésica Internacional como referencial geográfico pretende estudar a capacidade de recolha de informação de uma amostragem sistemática por linhas paralelas, de uma forma simples, expedita e repetível, relativamente à informação já disponibilizada pelas várias classificações do território de Trás-os-Montes e Alto Douro. A opção por um referencial geodésico implica uma densidade de amostragem crescente de Sul para Norte, em virtude da aproximação progressiva dos eixos de longitude entre o equador e os polos; a este facto contrapõe-se, em muito menor dimensão, o aumento da distância entre paralelos no mesmo sentido, devido à forma esférica e achatada do Planeta nos polos. Neste estudo, a distância entre os locais de amostragem mantém-se praticamente inalterada dentro dos eixos de longitude utilizados (aproximadamente 3062 metros), divergindo entre meridianos, de 20 812 metros à latitude 42°, até 21 118 metros à latitude 41°.

Do delineamento experimental descrito resultam 337 locais de amostragem; destes, houve necessidade de retirar 37, devido a limitações do *software* de análise utilizado. A selecção dos pontos de amostragem não considerados foi realizada pelo módulo HMCL do "package" SYNTAX IV (Podani, 1990), o qual identificou os pares de locais de amostragem de maior semelhança; de cada um dos 37 pares mais semelhantes rejeitou-se o elemento de maior número de ordem. O Quadro III.1 apresenta os pares de locais escolhidos e a quantificação da sua semelhança. No anexo 1 apresenta-se a localização aproximada dos locais de amostragem, representando-se por círculos a cheio os considerados na análise e por círculos vazios os não considerados. Não tendo este trabalho intenções de quantificação dos gradientes que investiga, conseguiu-se deste modo minimizar a influência de não se considerarem estes dados na análise dos resultados.

O número de 300 locais assume assim um compromisso entre os objectivos do estudo, a cartografia utilizada, e o tempo para realização do trabalho. O conhecimento da estrutura dos factores ecológicos da região, através da estratificação do território, necessita da consideração de todas as combinações possíveis dos mesmos. Ao considerar 300 locais de inventário, julgou-se abranger todos os complexos de atributos ecológicos do território. A cartografia utilizada, nas escalas 1/500 000 e 1/100 000, não aconselha o uso de maior precisão, o qual conduziria a falta de

QUADRO III.1 - Relação dos pontos retirados da análise, respectivo ponto mais semelhante e valor da sua dissemelhança.

Ponto retirado	Dissemelhança	Ponto mais semelhante	Ponto retirado	Dissemelhança	Ponto mais semelhante
45	0.00E+00	44	224	5.00E-02	223
48	0.00E+00	47	242	5.00E-02	241
154	0.00E+00	153	61	5.00E-02	60
212	0.00E+00	211	2	5.26E-02	1
270	0.00E+00	269	6	5.26E-02	5
314	0.00E+00	313	128	5.26E-02	127
163	2.44E-02	160	218	5.26E-02	217
320	2.56E-02	307	234	5.26E-02	233
153	3.03E-02	152	311	5.26E-02	296
156	4.76E-02	155	316	5.26E-02	313
284	4.76E-02	283	317	6.14E-02	313
332	4.76E-02	331	62	6.67E-02	60
162	4.84E-02	160	312	7.02E-02	296
24	5.00E-02	23	16	7.32E-02	15
26	5.00E-02	25	330	7.32E-02	322
107	5.00E-02	106	39	7.69E-02	38
138	5.00E-02	137	196	7.69E-02	194
239	5.00E-02	159	211	1.27E-01	210
208	5.00E-02	207			

rigor na recolha de dados.

3.2 APRESENTAÇÃO DOS FACTORES ESTUDADOS

"Classificação é abstracção. A escolha das propriedades usadas no diagnóstico ou diferenciação de características, não têm necessariamente de ser as mais importantes no seu sentido funcional. Elas são escolhidas em primeiro lugar pela sua facilidade de reconhecimento, de avaliação, etc. Elas não determinam necessariamente o carácter (a natureza) do objecto. Apenas o indicam." (Zonneveld, *in press*).

Neste estudo da ecologia do Nordeste de Portugal, não é realizada nenhuma avaliação prévia à funcionalidade dos factores, passível de "indicar" os sectores ambientais mais importantes. Esses factores podem ser dependentes e complexos, como os de ocupação do solo, ou primários e de menor complexidade, como a litologia, altimetria, etc... . A sua avaliação realiza-se *a posteriori*, sob a forma do contributo em termos de valor informativo para a classificação realizada. As variáveis possuem relações mais ou menos directas entre si; pretende-se com o método utilizado para o seu estudo, avaliar as redundâncias existentes entre elas. Foram por isso tratados de igual modo os factores primários ou de estação, e os factores secundários ou dependentes.

Os factores que controlam o desenvolvimento dos ecossistemas terrestres têm sido alvo de revisões; esses factores dependem, naturalmente, do ponto de vista de como é abordado o seu funcionamento (MOPU, 1992; Vos e Stortelder, 1992). As variáveis utilizadas neste estudo representam a informação que se considerou, de alguma forma, relevante para definição clara dos gradientes ecológicos da área. Elas repartem-se por três grandes grupos; o primeiro agrupa os atributos relativos ao ambiente climático, o segundo agrupa os atributos da base física do território e o terceiro agrupa os relativos à acção do Homem na utilização do espaço da província. Nos Quadros III.2 e III.3 apresentam-se as classes criadas para as variáveis, consoante a cartografia disponível seja do tipo isolinha ou do tipo área, respectivamente.

3.2.1 VARIÁVEIS CLIMÁTICAS

O papel do clima como factor ecológico é largamente reconhecido. Esse papel pode definir-se a diferentes níveis, os quais são geralmente designados por macroclima, mesoclima e microclima. Neste trabalho, as variáveis climáticas abordadas encontram-se entre o nível macro e mesoclimático, ou seja, entre o nível do clima regional e a sua influência por elementos da fisiografia do território. Essas variáveis são aqui abordadas sob o ponto de vista do seu desempenho no funcionamento ecológico do território, bem como da sua repartição geográfica.

QUADRO III.2 - Variáveis quantitativas utilizadas no estudo, com representação cartográfica do tipo isolinha.

Variável	Isolinha inferior	Isolinha superior	Amplitude de classe	Número de variáveis de presença/ ausência
Temperatura média anual	6	18	2	6
Temperatura máxima anual	12	24	2	6
Temperatura mínima anual	5	10	1	5
Amplitude térmica	5	17	2	5
Precipitação total anual	200	1400	200	6
Humidade relativa	40	100	20	3
Evaporação total anual	800	2000	200	6
Evapotranspiração total anual	500	900	100	4
Nº de dias de geada ¹	0	100	20	4
Nebulosidade	4	6	0.5	4
Insolação	450	650	50	4
Altitude	50	1150	100	11

A temperatura é um dos mais importantes factores climáticos. O sua acção faz-se sentir, tanto directamente sobre a fisiologia das plantas, como indirectamente sobre os processos geoquímicos, a hidrologia, a actividade biológica do solo, etc. Ao nível deste estudo, as temperaturas são influenciadas pela posição geográfica (litoral - interior) e pela topografia de cada local (altitude, exposição, etc.), interagindo ainda com outros factores como o vento, precipitação, humidade, etc. Tendo em conta a informação disponível, consideraram-se as temperaturas média (TMP.E_{__}), máxima (TMP.A_{__}) e mínima (TMP.I_{__}) médias anuais, a amplitude térmica anual (AMP.T_{__}), e o número médio de dias de geada no ano (GEA.N_{__}). Os valores das temperaturas e do número de dias de geada foram retirados da cartografia climática (Agroconsultores e COBA, 1990). A amplitude foi calculada pela diferença entre as temperaturas máximas e mínimas de cada local.

O padrão de temperaturas do nosso território segue os vales dos rios para os valores mais elevados, e as zonas serranas para os seus valores mais baixos. Os seus valores mais altos são observados na zona Sudeste, a área correspondente ao vale do rio Douro, quando este penetra o território nacional, deixando o seu curso internacional. Aí, as temperaturas atingem valores médios superiores a 16° C (TMP.E17) e máximos superiores a 22° C (TMP.A23). Os valores mais baixos ocorrem nos extremos Nordeste e Sudoeste (serra do Montesinho e serra do Montemuro), onde as mínimas descem abaixo dos 6° C (TMP.I55). O número de dias de geada (dias com temperaturas inferiores a 0° C) seguem aproximadamente o mesmo padrão, com os valores mais baixos (menos de 20 dias/ano - GEA.N10) a ocorrerem no percurso mais a jusante do rio Douro (Sudoeste do

¹ Os locais de amostragem não abrangem uma das classes extremas desta variável, pelo que o número de variáveis presença/ausência é menor do que o esperado atendendo ao valor das isolinhas inferior e superior e à amplitude considerada.

território), e os mais elevados (superiores a 80 dias/ano - GEA.N90) a ocorrerem nas maiores altitudes dos limites Norte e Sul da área de estudo.

A humidade atmosférica (vapor de água no ar) é também um importante factor ecológico, particularmente pelo seu papel na regulação do funcionamento dos mecanismos fotossintéticos das plantas. O seu valor está estritamente relacionado com a temperatura, nebulosidade, precipitação e quantidade de poeira na atmosfera. A humidade relativa representa a percentagem de vapor de água existente no ar, em relação ao máximo possível para determinadas condições de temperatura. No estudo considerou-se a humidade relativa média anual (HUM.R__). A cartografia da humidade relativa média anual da província apresenta 3 classes: valores superiores a 80% (HUM.R90), valores entre 60% e 80% (HUM.R70) e valores inferiores a 60% (HUM.R50). Os valores superiores a 80% (HUM.R90) ocorrem na fronteira Oeste da área, com um padrão semelhante aos maiores níveis altitudinais. Os valores mais baixos (HUM.R50) são característicos dos vales interiores dos afluentes do Douro (região central), bem como do seu curso internacional superior (fronteira Este da área de estudo). Toda a área restante da província apresenta valores intermédios (HUM.R70).

A precipitação é um factor ecológico que, embora com uma acção indirecta sobre os ecossistemas, tem muita importância na sua definição, na determinação dos usos do solo, e da paisagem em geral. A sua acção reflecte-se no ciclo hidrológico e humidade do ar durante o período vegetativo das plantas, bem como na temperatura do ar e do solo quando ocorre sob a forma de neve. Os registos da precipitação normalizados não diferenciam as diferentes formas da sua ocorrência. O seu valor representa a totalidade de precipitação, seja sob a forma líquida ou sólida (granizo e neve). Na cartografia dos valores da precipitação média anual (PRE.T__) utilizam-se classes de 200 l/m² de amplitude. Os valores mais baixos, inferiores a 400 l/m² (PRE.T03), ocorrem no canto Sudeste da área da província, no espaço correspondente ao vale do Douro, imediatamente antes e depois de deixar de ter curso internacional. Os valores mais elevados, superiores a 1200 l/m² (PRE.T13), acompanham os maiores níveis altitudinais dos limites Oeste (Barroso, Alvão, Marão e Montemuro) e Norte (Montesinho e Coroa) da província, bem como a serra da Nogueira, no seu interior Norte. Profundamente relacionada com a precipitação e outros elementos climáticos (humidade, temperatura, etc.), a nebulosidade expressa a fracção da esfera celeste que se apresenta coberta por nuvens. A presença de nuvens implica a diminuição da radiação, da transpiração das plantas e da evapotranspiração total dos sistemas, e o seu valor reflecte a proximidade ao litoral e a fisiografia local. A cartografia deste elemento (nebulosidade média anual) para a área da província representa-se por classes de 0.5 décimos (NUB.D__). O seu valor mais frequente situa-se entre 4.5 e 5.5 décimos (NUB.D53 e NUB.D48), atingindo os valores mais altos nas zonas montanhosas a Oeste, entre 5.5 e 6 décimos (NUB.D58), e os mais baixos na região mais interior do vale do Douro,

entre 4 e 4.5 décimos (NUB.D43). A insolação média anual é avaliada pelo número de horas de sol descoberto (Agroconsultores e COBA, 1990).

Os dados referentes às variáveis climáticas foram recolhidos de Agroconsultores e COBA (1990), no qual são representados por isolinhas à escala 1/500 000. A sua cartografia apoia-se em valores referentes ao período 1951/80, registados na rede de estações meteorológicas e postos udométricos normalizados da província e seus arredores. Todos os atributos climáticos cartografados são utilizados no estudo, exceptuando os baseados unicamente em valores de outros, já considerados directamente na análise. Estão neste caso vários índices e classificações climáticas (Thornthwaite e Köppen), regimes térmicos e hídricos do solo, etc.. A evaporação é avaliada pelo método de Piche, e a evapotranspiração potencial anual é avaliada pelo método de Thornthwaite (l/m^2).

As variáveis climáticas, apesar da sua natureza quantitativa, foram posteriormente transformadas em variáveis de presença/ausência por imperativos da metodologia utilizada no seu tratamento, sendo este aspecto desenvolvido adiante (ponto 3.3.1). Nos casos da amplitude térmica e número de dias de geada, não foi considerada a classe de menor valor (amplitude térmica entre 5 e 7° C e número de dias de geada entre 0 e 20), por nenhum destes casos ocorrer nos locais de amostragem. Foram deste modo constituídas 53 variáveis do tipo climático. No Anexo 3 g e h, e a título descritivo, são apresentados o mapa da temperatura média anual e o mapa da precipitação total anual para a região de Trás-os-Montes e Alto Douro.

QUADRO III.3 - Variáveis quantitativas e qualitativas utilizadas no estudo com representação cartográfica do tipo área.

Variável	Classes representadas	Número de variáveis de presença/ausência
Relevo	montanhoso,ondulado,aplanado	3
Declive	0% a 6%, 6% a 12/15%, 12/15% a 25/30%, 25/30% a 45/50%, > 45/50%	5
Geomorfologia	Vales largos, Superfícies de aplanamento, Resíduos quartzíticos	3
Afloramento Rochosos	10% a 25%, 25% a 50%, > 50%	3
Espessura do solo	< 10 cm, 10 cm a 50 cm, 50cm a 100 cm, > 100 cm	4
Muros de suporte	> 15/20 m, 8/10 m a 15/20 m, < 8/10 m	3
Litologia	granito, xisto, quartzitos, granodeoritos, migmitos, rochas básicas, rochas ultrabásicas, formações sedimentares, depósitos aluvionares	8
Solo ¹	leptossolos, fluvissoles, cambissolos, alissolos, luvissoles, podzóis, gleissolos, regossolos, antrossolos	10

¹ As litologias de depósitos aluvionares e os solos do tipo podzol, gleissolo e regossolo não se encontram representados nos locais de amostragem

3.2.2 VARIÁVEIS FISIAGRÁFICAS

Neste grupo de variáveis incluem-se os atributos relativos à base física do território em estudo. Caracterizou-se a altitude, o relevo, as particularidades geomorfológicas, a litologia e a pedologia. Para a caracterização da altitude recorreu-se à cartografia 1/100 000 do Instituto de Cartografia e Cadastro; optou-se por considerar classes de 100 metros de amplitude, entendida como a que melhor traduz esta variável para o tipo de trabalho e nível de precisão. Em relação ao relevo, consideram-se as classes utilizadas em Agroconsultores e COBA (1990); são elas os relevos montanhosos (superfícies muito onduladas ou acidentadas e com grandes declives), os relevos ondulados (superfícies suavemente onduladas com declives moderados), e os relevos planos (superfícies planas ou onduladas com pequenos declives). Também em relação ao declive foram consideradas as classes utilizadas em Agroconsultores e COBA (1990). Nos casos em que coincidem duas classes de declives, entendeu-se considerar a apresentada em primeiro lugar (a caracterização dominante). No que respeita à geomorfologia, consultou-se a "Carte Geomorphologique du Portugal" (Ferreira, 1981), considerando-se unicamente os elementos com representação espacial (vales fluviais bem marcados, superfícies de aplanamento e localizações de resíduos quartzíticos). Na caracterização dos afloramentos rochosos, ocorrência de muros de suporte e espessura do solo, também se recorreu à classificação de Agroconsultores e COBA (1990), tal como se apresenta no Quadro III.3.

Tanto os caracteres relativos à litologia como à pedologia têm origem na mesma fonte (Agroconsultores e COBA, 1990). Para a litologia, das classes representadas cartograficamente, apenas os depósitos aluvionares não ocorrem nos locais de amostragem. As unidades pedológicas consideradas por Agroconsultores e COBA (1990) são as definidas na legenda da Carta de Solos do Mundo da FAO-UNESCO. De entre as unidades consideradas, não ocorrem nos locais de amostragem as respeitantes a podzóis, gleissolos e regossolos. Devido à importância ecológica do factor solo, foram criadas variáveis de presença/ausência para referenciar a ocorrência de cada um dos seus tipos, tanto dominante como dominado. Deste modo, consideram-se as possíveis associações de tipos diferentes de solos, factor importante na definição de gradientes paisagísticos de um território.

Tal como para as variáveis climáticas, todas estas variáveis foram transformadas no tipo presença/ausência, perfazendo um total de 50. Não dispondo de outros elementos acessíveis do tipo fisiográfico, conseguiu-se uma caracterização satisfatória do substrato que serve de base ao desenvolvimento da paisagem. No Anexo 3 i e j, e a título descritivo, são apresentados o mapa geomorfológico e o mapa hipsométrico da região de Trás-os-Montes e Alto Douro.

3.2.3 VARIÁVEIS DE USO E OCUPAÇÃO DO SOLO

Para o reconhecimento dos principais usos e ocupações do território em estudo, recorreu-se à Carta dos Solos e Carta de Utilização Actual do Solo de Agroconsultores e COBA (1987), elaborada com base na interpretação de fotografias aéreas à escala 1/30 000 e transposta para cartas topográficas à escala 1/100 000. As classes consideradas são: sequeiro extensivo (Cs), sequeiro intensivo (Ch), regadio intensivo (Cr), olival (Po), vinha (Pv), amendoal (Pd), fruteiras (Pf), castanheiros (Pc), lameiros permanentemente húmidos (Lh), lameiros periodicamente húmidos (Ls), pastagem natural de altitude (La), culturas anuais e perenes circundando as povoações (Ei), matas e florestas de pinheiros (Mp), de carvalhos (Mc), de azinheiras, sobreiros e zimbreiras (Mz), e matos extremes ou com árvores dispersas (Im).

Para um estudo aprofundado da ecologia da Província, os elementos matas de pinheiros, matas de carvalhos e matas de sobreiros e azinheiras merecem uma maior discriminação. A diversidade das espécies de produção utilizadas nos repovoamentos florestais e o significado ecológico das diferentes espécies do género *Quercus* (caducifóleas, marcescentes e perenes) são fundamentais à compreensão da sectorização dos ambientes da área. Sendo exploradas de forma e com fins diversos, representam dinâmicas de utilização da paisagem bastante diferentes. Também a utilização de uma única classe de incultos, não discrimina a vegetação arbustiva que acompanha cada uma das ecologias percorridas pelo trabalho. Dever-se-ia ter em conta o valor indicativo das suas espécies e o seu estatuto na sucessão, factores importantes na interpretação da dinâmica da utilização do espaço (queimas, cortes, etc). No entanto, a informação disponível permite uma valiosa descrição relativa aos modelos de utilização do espaço pelo Homem, pelo que se optou pela sua utilização tal como publicada.

Tal como nas variáveis respeitantes ao solo, e pelas mesmas razões, também para os atributos de uso e ocupação do solo foram definidas variáveis de presença/ausência para os casos dominantes e para os dominados. A amostragem realizada e a frequência de ocorrência fez com que certas classes não fossem incluídas enquanto uso dominante ou dominado. Assim, não foi registada a ocorrência de regadio intensivo (Cr) e fruteiras (Pf) enquanto usos dominantes, e a ocorrência de pastagem natural de altitude (La) e culturas anuais e perenes circundando as povoações (Ei) enquanto usos dominados. A utilização dos restantes casos, transformados em variáveis presença/ausência, perfaz um total de 28 formas de uso e ocupação do solo. No Anexo 3 k e l, e a título descritivo, são apresentadas a distribuição das culturas arvenses de sequeiro e a de carvalhos (*Quercus robur* e *Quercus pyrenaica*) e castanheiros (*Castanea sativa*) da região de Trás-os-Montes e Alto Douro.

3.3 METODOLOGIA UTILIZADA

A estrutura da análise delineada para o estudo dos grandes grupos de ecologias da província de Trás-os-Montes e Alto Douro compreendeu metodologias de CLASSIFICAÇÃO, ANÁLISE DE CONCENTRAÇÃO e ANÁLISE DE INFORMAÇÃO. Com a CLASSIFICAÇÃO, pretendeu-se isolar sectores do território (grupos de locais) que, tendo identidade própria em termos dos respectivos complexos de factores ecológicos (grupo de atributos presentes), permitem considerá-los como a base sobre a qual se apoia a dinâmica paisagística da província. A ANÁLISE DE CONCENTRAÇÃO utiliza-se para conhecer a interligação entre os complexos de factores ecológicos e os sectores de território delineados. A ANÁLISE DE INFORMAÇÃO permite avaliar o desempenho de cada factor ecológico na organização estabelecida pelas análises anteriores.

3.3.1 CLASSIFICAÇÃO HIERÁRQUICA DIVISIVA

A teoria ecológica não pressupõe a existência de descontinuidades do seu objecto, ao contrário da taxonomia, em que a teoria da evolução apresenta a barreira reprodutiva como definidora da unidade básica individual que é a espécie. Assim, os métodos de classificação utilizados pelas duas disciplinas terão de reflectir estas diferenças; o taxonomista procura metodologias que lhe permitam unir as descontinuidades do seu objecto (métodos aglomerativos), enquanto o ecologista pretende metodologias que lhe permitam criar e sustentar descontinuidades no seu objecto, por forma a entender a sua estrutura (métodos divisivos).

A falta de recursos computacionais, bem como os resultados pouco credíveis que apresentavam, fizeram com que as técnicas divisivas de classificação fossem negligenciadas durante bastante tempo. Por outro lado, a taxonomia numérica desenvolvida pelas ciências naturais, sempre se baseou em técnicas aglomerativas, por razões inerentes à própria natureza do sujeito e tipo de investigação (estudo de organismos perfeitamente individualizados que necessitam do conhecimento das suas interligações ou semelhanças, não da sua fragmentação). O desenvolvimento da ecologia, e particularmente da fitossociologia nas suas disciplinas de sinfitossociologia e sintaxonomia, levaram a que se procurassem métodos que proporcionassem a divisão do que é observado como um aparente *continuum*, cujo padrão e funcionalidade lhe conferem uma individualidade própria.

Nas últimas três décadas, o uso de computadores e dos métodos numéricos a eles associados proporcionaram aos investigadores das ciências naturais ferramentas de análise de dados, entre as quais, as técnicas de classificação divisiva são as mais utilizadas pela ecologia. Estas técnicas permitem que populações, constituídas por um conjunto grande de indivíduos, sejam divididos em grupos mais pequenos, tendo como base a semelhança dos seus membros. São assim utilizadas para

a elaboração de tipologias, agrupamentos, partições, taxonomias numéricas, etc. (Everit, 1974; citado por Vos e Stortelder, 1992).

"TWINSpan" - Two Way INdicator SPecies ANalysis

Hill e col (1975) sumarizam 6 critérios que um método numérico deve cumprir para a classificação objectiva de dados de vegetação. Esses critérios, quando alargados a todos os dados de índole ecológica, descrevem-se como sendo os seguintes:

- (1) O método deve permitir classificar dados de aparência heterogénea, assim como providenciar uma conveniente ferramenta com a qual seja possível sumariar o intervalo de variação ecológica;
- (2) O método deverá ser de "resultado aberto", quer dizer, novos dados deverão ser facilmente localizados nos resultados anteriores, sem necessidade de repetição da totalidade dos cálculos;
- (3) O método deverá ser eficiente ao nível da computação, prático, e aplicável a amostragens em larga escala;
- (4) O método não deverá induzir erros graves por classificação duvidosa dos indivíduos;
- (5) O método deverá ser aplicável tanto a dados estacionais como de parcelas de ensaio;
- (6) Deverá ser possível utilizar um método de ordenação para extrair as maiores tendências dos dados, possibilitando uma visão geral da variação dos dados, providenciando uma base para a correlação com os inerentes factores ambientais responsáveis por essa variação.

Cumprindo tais pressupostos, Hill (1979) desenvolveu a técnica designada por "TWINSpan". "TWINSpan" é uma técnica de classificação do tipo divisivo, hierárquico e politético. Segundo Gauch (1982), a forma divisiva, para além das considerações tecidas antes, apresenta a vantagem de não necessitar da repetição dos cálculos quando surgem novos indivíduos, o que é inevitável quando se utilizam técnicas do tipo aglomerativo. Acresce ainda, a maior facilidade das técnicas divisivas em trabalhar grandes conjuntos de dados. A utilização em ecologia de técnicas do tipo hierárquico têm as vantagens de proporcionar análises individualizadas a diferentes níveis e exprimir as relações existentes entre os grupos constituídos. Ao considerarem vários caracteres como responsáveis pelas divisões (tipo politético), as classificações permitem avaliar o funcionamento comum desses caracteres, o que, tal como abordado em 2.1, revela grande significado ecológico (complexos de factores ecológicos).

A utilização do "TWINSpan" é recomendada por Gauch (1982) pela sua "robustez" e "efectividade"; de entre as suas vantagens, este autor destaca:

- a possibilidade de uso de dados originais;
- a classificação integrada de objectos e atributos;
- o rearranjo ordenado da matriz OBJECTOS x ATRIBUTOS inicial;
- a ordenação dos grupos, aproximando a apresentação dos mais similares;

- as necessidades computacionais relativamente baixas.

O programa "TWINSPAN" é convenientemente documentado no seu manual (Hill, 1979) e, para além de estudos de vegetação, temas tão diversos como a tipologia de solos florestais e o agrupamento de locais de nidificação de aves, foram estudados recorrendo à sua metodologia (Vos e Stortelder, 1992). Neste trabalho, esta técnica é utilizada para agrupar locais em função das variáveis climáticas, fisiográficas e de utilização e ocupação do solo que os caracterizam, tipificando assim os mais significativos sectores ecológicos do território em causa.

A estrutura do programa informático que permite a sua resolução é, de uma forma simplificada mas ao mesmo tempo esclarecedora, a seguinte (Hill, 1979):

1 - Estabelecimento de alguns parâmetros formais e operacionais; para além das indicações quanto ao formato de impressão utilizado, neste ponto são definidos limites e níveis para determinar:

- os atributos potencialmente preferenciais, em função da sua frequência nos dados;
- os atributos excluídos de, ou forçados a, ser indicadores, em função da preferência por um dos lados das dicotomias formadas;
- a frequência mínima para que um atributo seja ponderado;
- a aproximação mínima de um vector, para ser considerado vector próprio na análise de ponderação recíproca;
- o nível mínimo de preferência dos atributos por um dos lados da dicotomia, para ser ponderado na redefinição da ordenação, bem como o nível máximo para ser considerado como não preferencial;
- o número de repetições no refinamento da ordenação;
- a "desponderação" dos atributos raros, assim como a potenciação dos atributos preferenciais;
- o peso mínimo que é dado aos atributos;
- o intervalo entre o qual, as amostras são menos ponderadas no cálculo das preferências dos atributos por um dos lados da dicotomia;
- o tamanho da zona crítica para definição de classificações duvidosas;
- o número de segmentos da zona crítica, da zona de indiferença e de cada lado da zona crítica;
- os parâmetros de caracterização dos atributos, em função da sua fidelidade a cada uma das classes criadas para a sua classificação.

2 - Admissão duma matriz do tipo m objectos por n atributos; em formato fortran, complementado com a referência de cada atributo e de cada objecto.

3 - Definição pelo utilizador de parâmetros específicos, particulares do tipo de análise e/ou dados; para além dos estritamente formais, os que directamente vão condicionar o desenvolvimento da classificação:

- objectos que se pretende omitir;
- os níveis de abundância a considerar, consoante o tipo de dados;
- a dimensão mínima, em termos de número de objectos, de cada classe a formar;
- o número máximo de indicadores a considerar;
- o número máximo de atributos a ser apresentados no "output" final;
- o nível máximo de divisões requeridas;
- a apresentação ou não dos diagramas das divisões realizadas;
- a realização ou não de um ficheiro de resultados, próprio para ser lido por programas informáticos;
- a ponderação dos pseudo-atributos;
- os níveis de abundância, que sendo transformados em pseudo-atributos, são potenciais indicadores;
- os atributos que não devem ser considerados como indicadores.

4 - Conversão de níveis quantitativos em "pseudo-atributos".

5 - Classificação dos objectos.

6 - Classificação dos atributos em função do grau da sua preferência por cada grupo dos formados acima.

7 - Conversão das classificações de atributos e objectos em ordenações de modo a formar uma matriz diagonal positiva.

8 - Rearranjo da matriz de dados inicial de maneira que os atributos se apresentem na sua ordem correcta, definida em 7.

9 - Apresentação da tabela de dupla entrada com os resultados.

A estrutura do programa informático apenas permite o processamento de dados do tipo binário. São passíveis de serem utilizados dados qualitativos (ordinais e nominais), recorrendo à utilização de pseudo-atributos. Também os dados quantitativos podem, automaticamente e recorrendo a métodos aritméticos, ser transformados em binários, definindo-se pseudo-atributos por "níveis de corte", em função da ordem de grandeza do valor do atributo (quartis, decis, percentis, etc.). Este procedimento, se bem que vantajoso em numerosas situações, resulta pouco adequado na maior parte doutras. Vos e Stortelder (1992) apontam as seguintes desvantagens:

- apenas um conjunto de níveis de corte é admissível para todas as variáveis em jogo;
- é reconhecida a relevância de certos níveis, pelo que seria errado forçar outros limites artificiais;
- nem sempre os fenómenos ecológicos apresentam respostas lineares aos seus factores;
- os valores baixos são normalmente sobrestimados em relação aos mais elevados;
- os valores intermédios não são reconhecidos, sendo anulados pelo confronto entre os valores menores e maiores.

Neste estudo, tentando obviar tais problemas, optou-se por fazer manualmente a transformação de cada valor possível dos atributos numa variável presença/ausência. Tal foi facilitado porque os dados originais se encontravam já agrupados em classes, correspondendo estas às isolinhas da cartografia climática e altitudinal, ou a classes previamente determinadas (Agroconsultores e COBA, 1990).

A sua utilização neste trabalho deve-se sobretudo às boas referências que os mais variados autores fazem aos seus resultados e interpretação das suas soluções (Jongman e col, 1987). Os objectos são classificados pelo "TWINSPAN" em função das variáveis climáticas, fisiográficas e de utilização e ocupação do solo que os caracterizam. Dos resultados obtidos pela utilização do método "TWINSPAN", optou-se por considerar 8 classes de objectos e oito classes de atributos (terceiro nível de divisão). Como acima descrito, o número de classes a considerar corresponde a uma sucessão geométrica de base 2; considerou-se o valor de 4 como muito baixo, possivelmente indutor da perda de alguma divisão importante dos dados, e o valor de 16 muito elevado, não facilitando a interpretação dos resultados pela extrema estratificação ecológica que representa. O anexo 2 apresenta a tabela de dupla entrada produzida pelo programa, onde se distingue a estratificação de objectos e variáveis (respectivamente colunas e linhas), bem como a presença/ausência de cada atributo em cada local. Tal organização produz uma tendência para a apresentação de uma diagonal de valores de presença, o que, de uma forma não exclusiva, traduz a relação mais ou menos directa entre os grupos de objectos e de variáveis do estudo.

Classificação dos locais - grupos de objectos

Para além das variáveis indicadoras, foram definidas variáveis preferenciais, em função da tendência para a sua ocorrência num dos lados das dicotomias que se sucedem ao longo do processo divisivo. Tais divisões e seus resultados são comentados e apresentados em continuação.

A primeira divisão separa os 300 locais iniciais em dois grupos, de 220 (*grupo 0*) e 80 (*grupo 1*) elementos. Para isso utiliza como atributos indicadores, variáveis que permitem adivinhar a formação de um grande grupo de características continentais (maiores amplitudes térmicas e menor

nebulosidade), aparte de um grupo de locais com maior influência marítima, associada ou não à altitude (menores amplitudes térmicas e maior nebulosidade, e os maiores valores de precipitação).

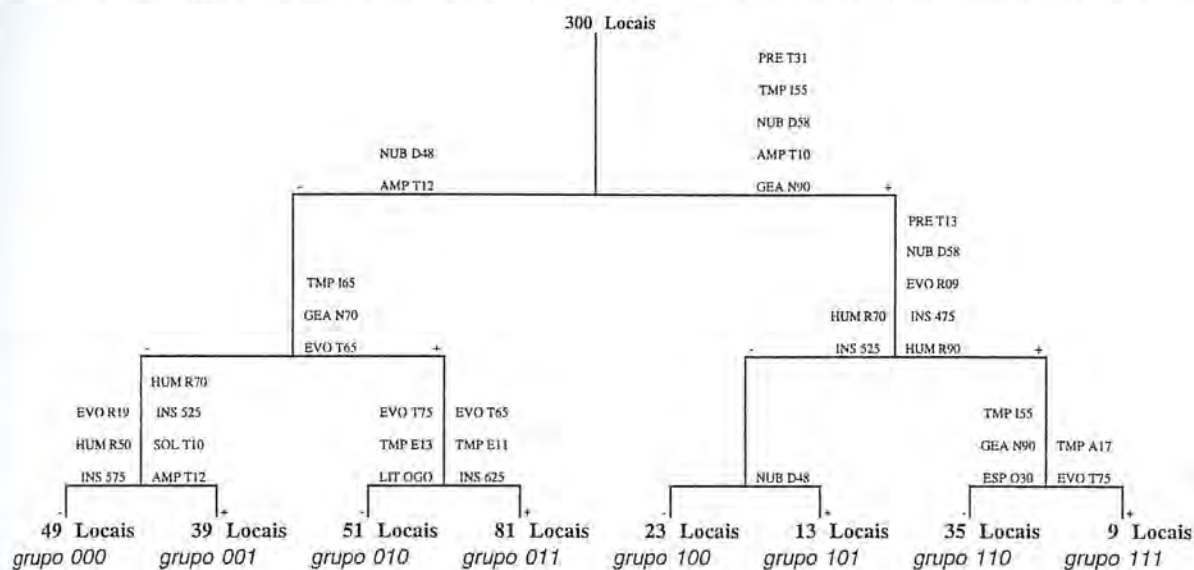


FIGURA 3.1 - Dendrograma relativo às partições realizadas pela metodologia do "TWINSPAN"

Se as variáveis indicadoras permitem conhecer as características dos grupos formados, as variáveis preferenciais permitem clarificar o significado dessa divisão (Quadro III.4). Entre estas, e para além de outras variáveis climáticas que seriam de adivinhar junto das indicadoras, são de grande significado as do tipo fisiográfico. Estão neste caso a altitude média de 700 metros (característica dos planaltos do interior transmontano) para o grupo 0, e os declives acentuados expondo afloramentos rochosos (serranias do limite Ocidental e Norte da província), para o grupo 1. Bastante sintomático das regiões em confronto são também os tipos de uso do solo dominante: o grupo 0 apresenta os cultivos extensivos de cereais e os olivais que dominam a paisagem continental; o grupo 1 prefere os pinhais, predominantes na paisagem serrana da zona marítima.

QUADRO III.4 - Variáveis preferenciais correspondentes à divisão entre os grupos 0 e 1.

Grupo de locais	Preferenciais climáticas	Preferenciais fisiográficas	Preferenciais de uso e ocupação do solo
<i>Grupo 0</i>	TMP.E13;TMP.E15;TMP.A19;TMP.I65; AMP.T12; AMP.T14; PRE.T07; PRE.T09; HUM.R50; EVO.R17; EVO.R19; EVO.T75; GEA.N50;NUB.D48; INS.575; INS.625	ALT.O07	UCS.O10; UPO.O05
<i>Grupo 1</i>	TMP.E09;TMP.E11;TMP.A13;TMP.A15; TMP.I55; AMP.T08; AMP.T10; PRE.T11; PRE.T13; HUM.R90; EVO.R09; EVO.R11; GEA.N90;NUB.D53; NUB.D58; INS.475; INS.525	DEC.O40; ROC.O20; SOL.B05	UMP.O10

Após a divisão inicial dos locais, passa-se a um segundo nível de divisão, onde cada um dos 2 grupos anteriormente referidos é, por sua vez, subdividido. Mais uma vez, as variáveis climáticas

assumem um papel de relevo na operacionalidade de tais divisões (variáveis indicadoras); no *grupo 0*, traduzem a distinção entre as regiões de clima mais ou menos favorável, dentro dos rigores do interior transmontano; no *grupo 1*, traduzem a distinção entre as zonas mais ou menos húmidas, na região de influência atlântica da província.

Os grupos resultantes da divisão do *grupo 0*, locais de influência continental, diferenciam-se climaticamente pela ameneidade do *grupo 00* (temperaturas mais altas, acompanhadas pelo menor número de dias de geada da área de estudo) e maior rigor do *grupo 01* (menores temperaturas e maiores precipitações e dias de geada). As variáveis fisiográficas ajudam a perceber que se está em presença da separação entre as zonas de vales interiores (altitudes baixas, grandes declives, geomorfologias de vale aberto, acompanhados da presença de solos antrópicos), e as zonas mais características do altiplano transmontano (altitudes medianas de relevos ondulados, cobertos por cambissolos). As variáveis preferenciais de uso e ocupação do solo indicam a separação de uma zona essencialmente de cultivos perenes (vinha e olival), de outra sem ocupação característica (Quadro III.5).

QUADRO III.5 - Variáveis preferenciais correspondentes à divisão entre os grupos 00 e 01.

Grupo de locais	Preferenciais climáticas	Preferenciais fisiográficas	Preferenciais de uso e ocupação do solo
<i>Grupo 00</i>	TMP.E15;TMP.A21;TMP.I75;TMP.I85; AMP.T14; PRE.T05; HUM.R50; EVO.R19; EVO.T75;EVO.T85; GEA.N30; GEA.N50; INS.525	ALT.O04; DEC.O40; DEC.O50; RIO.ABE; SOL.T10	UPO.O05; UPV.O10
<i>Grupo 01</i>	TMP.E11; TMP.A17; TMP.I65; PRE.T09; EVO.R15; EVO.R17;EVO.T65; GEA.N70; INS.625	ALT.O06; ALT.O07; REL.OND; SOL.B10	

A subdivisão do *grupo 1*, locais de maior influência atlântica, permite a distinção entre as zonas mais altas, e conseqüentemente mais húmidas, da região de influência atlântica da província. As variáveis preferenciais que elucidam tal divisão são sobretudo climáticas; no *grupo 10*, apenas a variável fisiográfica relativa às altitudes médias de 600 metros se intromete entre essa maioria do tipo climático; no *grupo 11*, onde aparecem os valores mais altos para a precipitação, humidade relativa, número de dias de geada e nebulosidade da área de estudo, aparecem as variáveis fisiográficas relativas às altitudes médias de 1000 metros e espessuras superficiais dos solos. Associadas a este quadro típico de montanha atlântica (*grupo 11*), surgem curiosamente as variáveis de uso e ocupação do solo relativas à presença de incultos na paisagem, como dominantes ou como dominados (Quadro III.6).

Os grupos do terceiro nível de divisão, definidos pela subdivisão dos 4 anteriores, permitem observar a consolidação de grupos de locais com características muito próprias, e grupos de transição entre esses mesmos. Eles resultam da operacionalidade de variáveis indicadoras, que na sua maioria, mas não em exclusividade como nos níveis anteriores, são do tipo climático. Essas

variáveis não climáticas são: a indicadora de solos antrópicos (*grupo 001* - isolando a região vinhateira do Douro), a indicadora de solos graníticos (*grupo 010* - separando as zonas com essa base litológica, das restantes constituintes dos altiplanos transmontanos), e a indicadora de grandes espessuras de solo (*grupo 110* - característicos das grandes "chãs" atlânticas que caracterizam os "plateaux" das elevações ocidentais mais importantes da província).

QUADRO III.6 - Variáveis preferenciais correspondentes à divisão entre os grupos 10 e 11.

Grupo de locais	Preferenciais climáticas	Preferenciais fisiográficas	Preferenciais de uso e ocupação do solo
<i>Grupo 10</i>	TMP.E13;TMP.I75; PRE.T09; PRE.T11; HUM.R70; EVO.R13; EVO.R15; EVO.T75; ; GEA.N70; NUB.D48; NUB.D53; INS.525	ALT.O06	
<i>Grupo 11</i>	TMP.E09; TMP.A13; AMP.T08; PRE.T13; HUM.R90; EVO.R09; EVO.T55; GEA.N90; NUB.D58; INS.475	ALT.O10; ESP.O20	UIM.O05; UIM.O10

A divisão do *grupo 00* isola dos restantes vales interiores da província (*grupo 000*), a região do Douro Vinhateiro (*grupo 001*). A acção do rio pode ser descrita como amenizadora dos rigores do clima, apresentando temperaturas máximas, amplitudes térmicas, evaporações e insolações mais baixas, e precipitações, humidades relativas e nebulosidades mais elevadas. Os solos antrópicos e esqueléticos, armados em patamares apertados, acolhem a cultura da vinha numa paisagem única na província e no País. Os restantes vales, não tão característicos, situam-se a altitudes superiores às do Douro, com relevos ondulados, cultivados predominantemente por arvenses de sequeiro (Quadro III.7).

QUADRO III.7 - Variáveis preferenciais correspondentes à divisão entre os grupos 000 e 001.

Grupo de locais	Preferenciais climáticas	Preferenciais fisiográficas	Preferenciais de uso e ocupação do solo
<i>Grupo 000</i>	TMP.A21; AMP.T14; PRE.T05; HUM.R50; EVO.R19; NUB.D43; INS.575	ALT.O03; REL.OND	UCS.O10
<i>Grupo 001</i>	TMP.E13; TMP.A19; AMP.T12; PRE.T09; PRE.T11; HUM.R70; EVO.R13; GEA.N30; INS.525	ESP.O10; SOL.T10; SOC.O30	UPV.O10

O *grupo 01*, ao subdividir-se nos *grupos 010* e *011*, sectoriza a região planáltica mais tradicional da província de Trás-os-Montes. O *grupo 010*, de clima mais ameno, diz respeito às regiões de origem granítica, com solos bem desenvolvidos, embora salpicados por afloramentos rochosos. A sua ocupação é dominada por florestas de pinheiro bravo, onde assume importância também, embora de forma dominada, a vinha (Quadro III.8). O *grupo 011* é a paisagem mais frequente e característica da província. De litologia básica e xistosa, e com declives medianos, o espaço é ocupado por grandes extensões cerealíferas, entrecortadas por leves depressões utilizadas para pastagem de regadio (lameiros) (Quadro III.8).

QUADRO III.8 - Variáveis preferenciais correspondentes à divisão entre os grupos 010 e 011.

Grupo de locais	Preferenciais climáticas	Preferenciais fisiográficas	Preferenciais de uso e ocupação do solo
<i>Grupo 010</i>	TMP.E13; TMP.A19; AMP.T14; EVO.R13; EVO.T75	ROC.O20; ESP.O20; LIT.OGO; SOL.I05; SOL.B10	UPV.O05; UMP.O10
<i>Grupo 011</i>	TMP.E11; EVO.T65; INS.625	DEC.O20; LIT.OXO; LIT.OBO	UCS.O10; ULS.O05

O grupo 10 corresponde a uma região "média atlântica" e divide-se no grupo 100, de clima mais frio e fisiografia agreste de origem xistosa, e no grupo 101, de clima e fisiografia mais brandos, que embora de solos esqueléticos e frequentes afloramentos rochosos, são explorados eficientemente graças à sua organização em largos patamares. As regiões do grupo 100 não apresentam uma ocupação característica dominante, sendo exploradas as culturas de sequeiro extensivo (cereais de inverno) de forma dominada. Já a paisagem do grupo 101 é dominada por culturas de sequeiro intensivo, que ocupam os largos socialcos entrecortados por pequenos bosquetes de pinheiros (Quadro III.9).

QUADRO III.9 - Variáveis preferenciais correspondentes à divisão entre os grupos 100 e 101.

Grupo de locais	Preferenciais climáticas	Preferenciais fisiográficas	Preferenciais de uso e ocupação do solo
<i>Grupo 100</i>	TMP.E13; TMP.A17; TMP.I65; TMP.I75; AMP.T08; EVO.R13; NUB.D53; INS.575	REL.MON; DEC.O40; ROC.O20; LIT.OXO	UCS.O05
<i>Grupo 101</i>	TMP.A15; TMP.I55; PRE.T09; EVO.R15; NUB.D48	ALT.O05; ALT.O06; REL.OND; DEC.O20; ROC.O30; ESP.O10; SOL.T10; SOC.O10	UCH.O05; UCH.O10; UMP.O05

O grupo 11, o grupo das regiões de maior altitude e influência atlântica, divide-se no grupo 110, que diz respeito às cumeadas, e no grupo 111, respeitante às zonas de meia encosta. O grupo 110, naturalmente de clima mais frio, caracteriza-se por grandes espaços abertos e planos, de solos desenvolvidos derivados de granito e pontilhados por afloramentos rochosos. Sem uma ocupação característica, são áreas cultivadas, entrecortadas por incultos. Bastante características são as alturas do Barroso e as chãs do Alvão, com a sua especialização na produção de batata de semente, e o planalto de Montesinho, ocupado na sua maior parte por pastagens semeadas. O grupo 111, o mais pequeno da classificação, caracteriza as regiões de meia encosta, de clima muito mais favorável, onde o Homem se instalou, construindo socialcos de largura mediana, onde faz as culturas anuais e perenes que circundam as suas povoações (Quadro III.10).

QUADRO III.10 - Variáveis preferenciais correspondentes à divisão entre os grupos 110 e 111.

Grupo de locais	Preferenciais climáticas	Preferenciais fisiográficas	Preferenciais de uso e ocupação do solo
<i>Grupo 110</i>	TMP.E09; TMP.E11; TMP.A13; TMP.A15; TMP.I55; AMP.T08; HUM.R70; EVO.T55; EVO.T65; GEA.N90; NUB.D53; INS.575	ALT.O09; ALT.O10; ALT.O11; REL.OND; DEC.O20; SUP.APL; ROC.O20; ESP.O30; LIT.OGO	UIM.O05
<i>Grupo 111</i>	TMP.E13; TMP.E15; TMP.A17; TMP.I75; TMP.I85; AMP.T12; EVO.R11; EVO.T75; GEA.N50	ALT.O05; ESP.O10; ESP.O20; LIT.OXO; SOL.T10; SOC.O20	UELO10

Classificação dos atributos

Paralelamente à classificação dos locais em estudo, também as variáveis utilizadas são divididas em igual número de classes. Tal partição é directamente resultante da divisão antes efectuada para os locais, reflectindo a coincidência dos atributos nos grupos de locais previamente definidos. No Quadro III.11 apresenta-se a repartição final das variáveis utilizadas no estudo.

QUADRO III.11 - Classificação dos atributos utilizados no estudo.

Grupo 000	Grupo 001	Grupo 010	Grupo 011	Grupo 100	Grupo 101	Grupo 110	Grupo 111
TMP E15	TMP E13	TMP I75	TMP A17	PRE T11	TMP E11	TMP E07	ROC O20
TMP E17	TMP I65	HUM R70	EVO R15	EVO R13	NUB D53	TMP E09	
TMP A19	AMP T12	ALT OO5	EVO T65	INS 525	ALT OO8	TMP A13	
TMP A21	PRE T09	REL MON	GEA N70	DEC O40	ROC O30	TMP A15	
TMP A23	EVO R17	REL OND	ALT OO6	LIT OGO	SOL B05	TMP I55	
TMP I85	NUB D48	DEC O20	UCS O05	SOC O10	UPC O10	AMP T08	
TMP I95	INS 575	DEC O30	UCH O05	SOC O20	UMC O05	AMP T10	
AMP T14	INS 625	SUP APL	ULS O10	UEI O10	UMC O10	PRE T13	
AMP T16	ALT OO7	ESP O10	UMP O05	UMP O10		HUM R90	
PRE T03	REL PLA	ESP O20				EVO R09	
PRE T05	DEC O10	ESP O30				EVO R11	
PRE T07	RES QUA	LIT OXO				EVO T55	
HUM R50	LIT ODO	SOL I10				GEA N90	
EVO R19	LIT OMO	SOL B10				NUB D58	
EVO T75	LIT OBO	SOL T05				INS 475	
EVO T85	LIT OOU	SOL T10				ALT OO9	
GEA N30	LIT OSO	UCH O10				ALT O10	
GEA N50	SOL I05	UCR O05				ALT O11	
NUB D43	SOL U05	UPF O05				ULH O05	
ALT OO1	SOL U10	UIM O05				ULH O10	
ALT OO2	SOL L05	UIM O10				ULA O10	
ALT OO3	SOL L10						
ALT OO4	UCS O10						
DEC O50	UPV O05						
RIO ABE	UPC O05						
ROC O10	ULS O05						
ESP O40							
LIT OQO							
SOC O30							
UPO O05							
UPO O10							
UPV O10							
UPD O05							
UPD O10							
UMZ O05							
UMZ O10							

De dimensão muito variável, a repartição das variáveis segue, em traços largos, a caracterização dos locais dos grupos com a mesma referência. Os gradientes caracterizados, dos mais continentais para os mais marítimos, ficam assim traduzidos pelas variáveis que mais frequentemente contribuem para a sua caracterização. Para tal, como referido na metodologia do "TWINSPAN", concorre o facto de ser com base na repartição dos locais que o programa faz a atribuição das diferentes variáveis a cada um dos grupos, permitindo a construção de uma tabela de dupla entrada, com a sua diagonal principal preenchida por valores de presença. Tal é também o aspecto do Quadro III.12 onde se apresenta a quantidade de "presenças" para cada par grupo de locais/grupo de atributos.

QUADRO III.12 - Tabela de contingência entre grupos de locais e grupos de variáveis definidos.

		<i>Grupo de Locais</i>							
		<i>000</i>	<i>001</i>	<i>010</i>	<i>011</i>	<i>100</i>	<i>101</i>	<i>110</i>	<i>111</i>
Grupo de Variáveis	000	486	288	171	145	18	9	3	22
	001	154	129	310	559	40	25	34	13
	010	289	255	315	487	138	81	185	52
	011	12	233	126	298	64	38	34	15
	100	28	85	82	23	76	40	54	16
	101	10	9	24	101	52	29	50	7
	110	5	20	11	2	58	47	325	50
	111	6	4	17	1	9	1	15	0

3.3.2 ANÁLISE DE CONCENTRAÇÃO

A análise canônica da tabela de contingência relativa às frequências dos grupos de locais/grupos de atributos (Quadro III.12), permite avaliar a responsabilidade desses grupos de atributos na definição da ecologia de cada grupo de locais definido. Mais do que o valor dos seus coeficientes, a sua visualização num espaço bidimensional permite compreender as tendências dos grupos de locais e de atributos.

A utilização do programa CONAPACK (Orlóci, 1991b) permite a definição de 5 eixos canônicos, com os resultados apresentados no Quadro III.13.

QUADRO III.13 - Teste de χ^2 e coeficientes de correlação canônica entre grupos de locais e grupos de atributos.

	Coeficiente de correlação canônica	Valor de χ^2 associado a cada eixo	Valor de χ^2 acumulado
Primeiro eixo	0.597	2238	62%
Segundo eixo	0.376	890	87%
Terceiro eixo	0.213	284	95%
Quarto eixo	0.144	130	98%
Quinto eixo	0.096	58	100%
Total		3600 (49 g. l.)	

Na figura 3.2 representam-se os grupos de locais e grupos de variáveis em relação ao espaço definido pelos dois primeiros eixos canônicos (87 % do total de χ^2). Os grupos de locais são representados em itálico, e os grupos de atributos em negrito. No referido gráfico, o primeiro eixo (horizontal) apresenta-se como separador entre os grupos 0 e os grupos 1 (ecologias continentais à esquerda vs oceânicas à direita). O segundo eixo (vertical) aparece como diferenciador dos grupos 0, aqui com sentido inverso ao definido pelo "TWINSPAN". Os grupos 000 e 011 apresentam uma quase perfeita sobreposição entre os locais e os seus atributos caracterizadores, o que não acontece com os outros grupos, particularmente os oceânicos (grupos 1). Entre estes, é de notar a proximidade evidenciada entre os grupos de locais *100* e *101*, o qual não tem expressão nos mesmos

grupos de atributos. Os grupos de factores **010** e **100** são os que mais são compartilhados pelos grupos de locais (centro da representação), apresentando-se os restantes como diferenciadores (extremos da representação), casos de **000** e **110**.

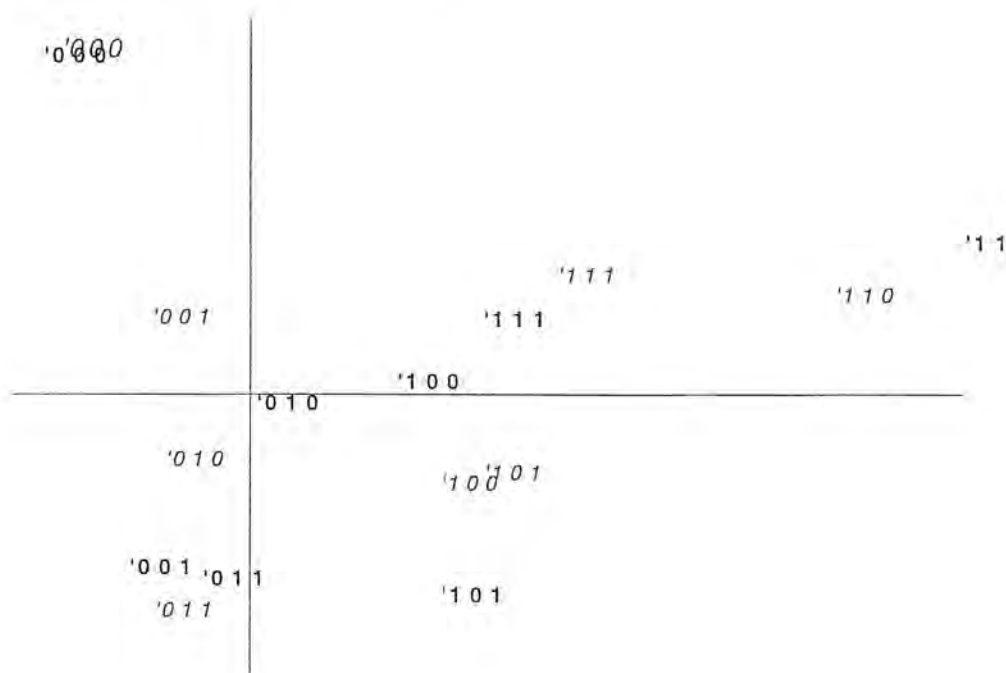


FIGURA 3.2 - Representação gráfica dos grupos de locais e grupos de atributos no espaço definido pelos dois primeiros eixos canônicos .

3.3.3 ANÁLISE PELA TEORIA DA INFORMAÇÃO

O conhecimento das inter-relações entre as variáveis ecológicas é um das mais importantes componentes da ciência ecológica (Lepart e Debussche, 1980). Os atributos caracterizadores de um território podem ser avaliados em função da sua "importância" na definição da ecologia desse território. Em sentido restrito, esta "importância" diz respeito ao sucesso de cada variável na participação na variação comum ao conjunto da amostragem realizada. Dessa forma conhecem-se as variáveis que mais descrevem os gradientes ecológicos do território, podendo desse modo dispensar a avaliação de outras com menos sucesso nessa indicação (Orlóci, 1976).

As relações entre variáveis são normalmente avaliadas por parâmetros como a covariância. No entanto, para variáveis binárias, tais parâmetros não são os mais indicados (Orlóci, 1976). Os dados binários traduzem a presença ou ausência de determinado carácter, não existindo uma ordem de grandeza que interfira na quantidade comum de variação. A eficiência informativa atribuída a cada

variável pode ser avaliada pela sua média de informação comum com as outras variáveis, ou ainda, pelo ganho do conjunto das variáveis estudadas nessa mesma quantidade, pelo facto de coexistirem.

QUADRO III.14 - Ordenação dos atributos pelo seu valor em informação mútua.

Atributos climáticos			Atributos fisiográficos			Atributos de uso do solo		
Variáveis	Nº de ordem	Informação mútua	Variáveis	Nº de ordem	Informação mútua	Variáveis	Nº de ordem	Informação mútua
PRE T13	1	1334.226	SOL I10	5	968.2026	UCS O10	33	358.0799
EVO T65	2	1200.143	REL MON	8	878.3099	UIM O10	43	272.0804
NUB D58	3	1058.089	ESP O30	15	673.908	UPV O10	54	173.7866
TMP I65	4	1033.632	LIT OGO	20	563.5004	UMP O10	64	128.2591
TMP E11	6	909.8101	REL OND	24	474.7235	UIM O05	66	119.7123
GEA N90	7	898.8812	SOL B10	25	468.0629	UPV O05	68	104.6683
EVO R19	9	847.9517	DEC O30	36	317.4842	UPO O05	69	95.86224
TMP A17	10	828.9353	SOL T10	37	314.994	UMZ O05	72	87.74487
TMP I55	11	776.472	ESP O20	40	289.9715	UCH O05	76	73.1479
TMP E15	12	756.8403	LIT OXO	44	258.4456	UPD O05	78	69.12749
NUB D48	13	744.0569	ALT O07	46	241.0791	UCS O05	81	64.3579
EVO R09	14	693.3511	DEC O50	50	209.8352	ULS O05	88	45.34841
INS 625	16	666.425	DEC O10	51	193.7197	UCH O10	93	32.20528
HUM R50	17	658.4664	DEC O40	53	175.1806	UPF O05	95	29.03013
GEA N70	18	625.5753	RIO ABE	56	167.0281	UMP O05	96	25.86758

O algoritmo usado para ordenar as variáveis usadas neste trabalho foi desenvolvido por Podani (1990) no módulo SUMR, do seu "package" SYNTAX IV. O método utilizado elimina o efeito dos atributos à medida que estes são seleccionados, ordenando as variáveis na combinação que transmite a máxima informação do conjunto dos dados. O programa selecciona a variável para a qual a soma de informação mútua com todas as outras variáveis é máxima, não sendo o seu contributo para a totalidade da informação mútua entre as variáveis, contabilizado nos cálculos seguintes. Os resultados apresentam-se no Quadro III.14, onde se apresentam as 15 primeiras variáveis de cada tipo considerado, caracterizadas pela respectiva informação mútua e número de ordem no "ranking" das 131 estudadas.

Dos resultados obtidos retém-se a ideia da maior importância das características climáticas, seguidas das fisiográficas e das de uso do solo, na definição dos seus gradientes. Os valores de número de ordem e de informação mútua revelam a grande disparidade existente entre o valor informativo dos atributos dos três tipos. Tal ideia já tinha sido expressa por van de Maarel (1976; citado por Lepart e Debussche, 1980); este autor considera as variáveis climáticas como as mais importantes ao nível da região ecológica, sendo as litológicas e as topográficas mais significativas ao nível do sector ecológico. Esta ideia foi inicialmente apresentada por Manil (1963; citado por Lepart e Debussche, 1980), concluindo que as relações entre variáveis ecológicas dependem grandemente do tamanho da área em estudo. O mesmo autor refere ainda o tipo de área em estudo como determinante da importância das variáveis na definição dos gradientes ecológicos. Essa é

também a opinião de Rougerie e Beroutchachvili (1991), que comentam serem as variáveis climáticas, determinantes dos *climax* em territórios acidentados, como é o de Trás-os-Montes e Alto Douro.

A capacidade integradora de uma classificação, sintetizando a evolução dum conjunto de variáveis num território, pode ser avaliada pela coincidência entre essas variáveis e os sectores definidos pela classificação. Quanto maior fôr essa coincidência, maior será a probabilidade da presença dessa variável significar a presença de determinado sector. De Pablo e col (1987), num estudo sobre a região de Madrid, abordam a relação entre os conceitos de incerteza e informação e sua utilização na avaliação, quer da capacidade integradora de determinada classificação, quer da eficácia descritiva das variáveis consideradas. A informação total ou entropia total de um sistema composto por uma variável e um sector da classificação define-se por:

$$H(V.S) = H(V) + H(S|V) = H(S) + H(V|S)$$

com $H(V)$ - incerteza associada à presença de uma variável num ponto do território,

$H(S|V)$ - incerteza associada à participação de um ponto do território num sector da classificação, tendo em conta a presença ou ausência de determinada variável nesse local,

$H(S)$ - incerteza associada à participação de um ponto do território num sector de classificação e

$H(V|S)$ - incerteza associada à presença ou ausência de uma variável num ponto do território, quando é conhecido o respectivo sector da classificação.

A coincidência entre os sectores definidos por uma classificação e as variáveis que lhe estão na base indica o valor informativo dessas variáveis. A informação que cada variável possui sobre a sectorização do território é quantificada por (De Pablo e col, 1987):

$$I(V.S) = H(V) - H(V|S) = H(S) - H(S|V) = H(S) + H(V) - (H(V.S)).$$

Os mesmos autores propõe a expressão

$$C = \frac{I(V.S)}{H(V)}$$

para avaliar o valor predictivo de uma variável sobre um conjunto de sectores definidos por uma classificação.

No Quadro III.15 apresentam-se os valores de *C* mais elevados, obtidos para os três grupos de variáveis utilizados no estudo. Mais uma vez se destaca o papel preponderante das variáveis climáticas na indicação dos sectores ecológicos definidos para Trás-os-Montes e Alto Douro (valores de *C* mais elevados), secundadas pelas variáveis fisiográficas e pelas de uso do solo, estas as de menor valor. Das variáveis fisiográficas, as de maior valor informativo são as relativas à altitude e presença de socacos. Em relação às variáveis de uso do solo, destaca-se o valor informativo das relativas a pastos permanentes e culturas perenes, tanto como uso dominante como uso dominado.

QUADRO III.15 - Valores de "*C*" mais elevados para os atributos climáticos, fisiográficos e de uso do solo, na classificação realizada com "TWINSPAN".

Variáveis climáticas		Variáveis fisiográficas		Variáveis de uso do solo	
PRE T13	0.848	SOC O10	0.478	ULA O10	0.391
NUB D58	0.738	SOL T10	0.475	ULH O10	0.362
HUM R90	0.69	ALT O10	0.471	ULH O05	0.33
TMP I55	0.655	ALT O11	0.433	UPD O10	0.328
TMP A15	0.633	SOC O30	0.393	UPV O10	0.273
EVO T65	0.633	ALT O03	0.343	UPC O05	0.238
GEA N90	0.626	SOC O20	0.336	UPF O05	0.237
EVO R09	0.573	ALT O09	0.321	UMP O10	0.235
TMP E09	0.568	ALT O01	0.301	UCR O05	0.23
EVO R19	0.531	ESP O10	0.292	ULS O05	0.208

Os valores mais elevados de *C* correspondem a variáveis que integram os grupos que mais se destacam nos dois primeiros eixos canónicos da figura 3.2, o que confirma o seu valor discriminante nos gradientes da província. O grupo de variáveis **010** agrupa a maioria das variáveis climáticas do "ranking" do Quadro III.15, bem como parte das fisiográficas e de uso do solo mais associadas aos ambientes atlânticos da província. O grupo de variáveis **000** agrupa grande parte das variáveis fisiográficas e de uso do solo do "ranking" do mesmo Quadro, indicadores dos gradientes de clima mediterrâneo da província. Em relação às variáveis preferenciais dos grupos de locais, as variáveis com valores mais importantes de *C* do Quadro III.15 correspondem na sua maioria, aos grupos **000**, **001**, **110** e **111**, ou seja, aos grupos extremos de gradientes de clima mediterrâneo e atlântico da província. A sua distribuição confirma a incidência desses atributos nos grupos de variáveis citados antes, assim como a importância relativa das variáveis climáticas para distinguir os sectores "atlânticos" e das variáveis fisiográficas para distinguir os sectores "mediterrâneos".

3.4 COMPARAÇÃO DOS RESULTADOS COM OUTRAS CLASSIFICAÇÕES

Os resultados obtidos foram contrastados com outras classificações existentes, quer de âmbito regional, quer de âmbito nacional. Pretendeu-se assim "localizar" os resultados, em trabalhos já realizados no mesmo âmbito. Essas classificações anteriores são de base empírica e, embora apoiadas em dados da rede climatológica nacional, traduzem sobretudo a muito valiosa experiência e conhecimento de "campo" dos seus autores. Elaboradas com fins diversos, elas possuem

características que permitem ser comparadas com os resultados agora obtidos; para além do seu âmbito ecológico, têm expressão cartográfica, o que permite que sejam contrastadas com o esquema de amostragem utilizado neste trabalho.

3.4.1 GRANDES REGIÕES (AGROCONSULTORES E COBA, 1990)

A carta das Grandes Regiões foi publicada na 2ª versão preliminar da Carta de Solos, Carta da Utilização Actual do Solo e Carta do Uso Potencial do Solo do Nordeste de Portugal (Agroconsultores e COBA, 1990). Com objectivos meramente operacionais, visava uma compartimentação do território que facilitasse a elaboração da Carta de Solos. Os seus autores justificam tal facto com a grande variabilidade de factores como o clima, a litologia e a geomorfologia, na área geográfica em causa; deste modo, seria conseguida uma sistematização dos estudos a realizar (Anexo 4 n).

"Grandes Regiões", segundo a definição de Agroconsultores e COBA (1990) são grandes espaços geograficamente contínuos e com características climáticas e morfológicas capazes de as distinguir umas das outras. Esta combinação confere-lhes a identidade capaz de as individualizar. Tal divisão do território é o nível mais elevado numa estratificação que é complementada por sub-regiões e zonas homogéneas. As sub-regiões são delimitadas em função ainda de aspectos morfológicos e climáticos particulares, de âmbito mais restrito. As zonas homogéneas têm em conta a base litológica do território.

Surgem assim 13 Grandes Regiões, as quais podem ser divididas em (Agroconsultores e COBA, 1990):

- Regiões com grande dominância de terras altas (superiores a 650/700 metros) em planaltos e montanhas, de influência marcadamente atlântica - BARROSO, ALVÃO-MARÃO, PADRELA, MONTEMURO e BEIRA DOURO - ou marcadamente continental - BRAGANÇA, CARRAZEDA e MIRANDA-MOGADOURO.

- Regiões com grande dominância de terras de altitudes médias (entre 450/500 metros e 650/700 metros) - BORNES-SABOR.

- Regiões com grande dominância de terras baixas (menores de 450/500 metros) com maior influência atlântica - TÂMEGA - ou mais influência continental - DOURO, TUA e ALTO DOURO.

Estas divisões, mais do que propriamente classificar, visam uma repartição do território. Tal facto não lhe retira no entanto, aspectos de integração de diversos factores ecológicos do território, pelo que se achou importante a sua confrontação com os resultados agora obtidos. A sua base

empírica denota a determinação de uma hierarquia dos macrofactores que contribuem para a definição ecológica do território. A decisão pela escolha de macrofactores climáticos e morfológicos revela-se fundamental, o que de certa forma é confirmado pelos resultados agora obtidos.

Para a sua comparação com os resultados obtidos pelo "TWINSPAN", foi identificada a Grande Região que corresponde a cada local de amostragem deste estudo. Recorreu-se à análise canónica da tabela de frequências cruzadas com os grupos de locais classificados pelo "TWINSPAN" (Quadro III.16), num processo análogo ao realizado no ponto 3.3.2, para avaliar a coincidência entre os grupos de locais e os grupos de atributos.

QUADRO III.16 - Tabela de contingência entre os grupos de locais definidos pelo "TWINSPAN" e as classes da classificação das Grandes Regiões.

Grupos	000	001	010	011	100	101	110	111	Total
Barroso	0	0	0	0	0	0	19	3	22
Montemuro	0	3	1	0	0	0	2	0	6
Tâmega	0	0	6	0	12	0	0	5	23
Alvão-Marão	0	0	0	0	1	0	7	1	9
Douro	2	27	3	0	0	0	0	0	32
Beira Douro	0	5	8	0	0	13	0	0	26
Padrela	0	0	16	0	3	0	1	0	20
Bragança	0	0	4	25	4	0	6	0	39
Tua	14	0	3	7	0	0	0	0	24
Carrazeda	4	0	2	0	0	0	0	0	6
Alto Douro	20	4	6	4	0	0	0	0	34
Bornes-Sabor	3	0	2	14	3	0	0	0	22
Miranda-Mogadouro	6	0	0	28	0	0	0	0	34
Total	49	39	51	81	23	13	35	9	300

No Quadro III.17 apresentam-se os dados referentes aos 6 eixos canónicos obtidos a partir da tabela de contingência referente à comparação entre as classificações do "TWINSPAN" e as Grandes Regiões (Agroconsultores e COBA, 1990). Nela se pode constatar a coincidência bastante significativa entre as duas classificações, atestada pelo valor de χ^2 . Os coeficientes de correlação, elevados para os dois primeiros eixos, confirmam tais resultados.

A figura 3.3 representa a localização relativa dos grupos das duas classificações, em relação ao espaço definido pelos dois primeiros eixos canónicos. O primeiro eixo revela as diferenças entre os extremos atlânticos e continentais da província, e o segundo, as diferenças existentes entre as regiões de "transição". É notável a sobreposição do grupo 110 com as Grandes Regiões Barroso e Alvão-Marão (Oeste Atlântico) e do grupo 001 com a região do Douro (Este Mediterrânico). Em relação aos últimos grupos, a classificação de Grandes Regiões distingue a região do Douro (vinhateiro) da região Alto-Douro (internacional), o que leva à aproximação da primeira ao grupo definido por 001 e da segunda ao grupo definido por 000.

QUADRO III.17 - Teste de χ^2 e coeficientes de correlação canónica entre grupos de locais definidos pelo "TWINSpan" e a classificação das Grandes Regiões.

	Coefficiente de correlação canónica	Valor de χ^2 associado a cada eixo	Valor de χ^2 acumulado
Primeiro eixo	0.865	222	26%
Segundo eixo	0.835	207	50%
Terceiro eixo	0.716	152	68%
Quarto eixo	0.643	123	83%
Quinto eixo	0.583	101	94%
Sexto eixo	0.393	46	100%
Total		851 (84 g.l.)	

A confrontação das duas classificações põe em relevo que a separação de Grandes Regiões como Barroso e Alvão-Marão se deve a motivos operacionais, ou seja, pela sua distância física. Fica patente a sua identidade ecológica, pelo menos ao nível de discriminação realizado neste estudo. A semelhança entre as regiões de Douro e Beira Douro e os correspondentes grupos 001 e 101 revelam o carácter integrador proporcionado pelo elemento fluvial fundamental da província.

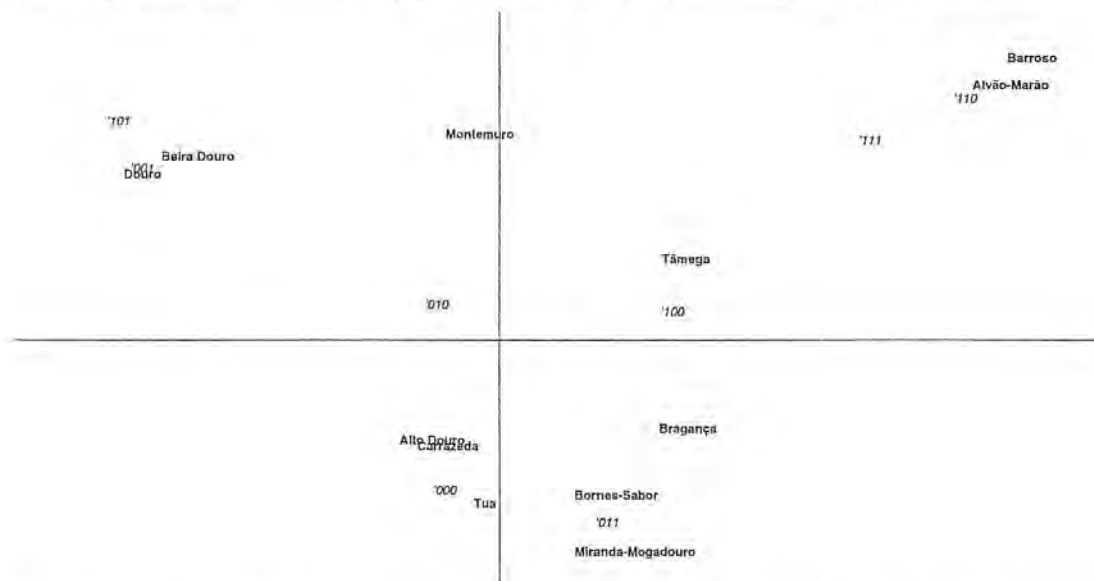


FIGURA 3.3 - Representação gráfica dos grupos de locais definidos pelo "TWINSpan" e as classes da classificação das Grandes Regiões (Agroconsultores e COBA, 1990), no espaço definido pelos dois primeiros eixos canónicos.

Para avaliar o papel que desempenham os atributos na indicação das Grandes Regiões, foram calculados os valores de C , de forma igual à realizada para a classificação do "TWINSpan". A observação dos resultados (Quadro III.18) revela que, também nesta classificação, a importância dos factores climáticos se sobrepõe aos fisiográficos e aos de uso do solo, embora com valores inferiores aos da classificação anterior ("TWINSpan"). Por outro lado, os valores para os atributos fisiográficos e de uso do solo são muito semelhantes, sendo os cinco mais elevados de uso do solo, superiores aos cinco mais elevados fisiográficos, o que não acontecia antes ("TWINSpan"). Para tal facto deverá concorrer a própria natureza da classificação das Grandes Regiões; procurou-se realizar uma zonagem de *continua*, que apresentasse constância dos factores mais importantes para a formação do solo, o que naturalmente dará um pouco mais de importância aos fisiográficos e de uso

do solo. No entanto, fica mais uma vez patente a notabilidade da importância do clima quando se pretende abordar o território de Trás-os-Montes e Alto Douro.

QUADRO III.18 - Valores de "C" mais elevados para os atributos climáticos, fisiográficos e de uso do solo, na sectorização realizada pela classificação das Grandes Regiões.

Variáveis climáticas		Variáveis fisiográficas		Variáveis de uso do solo	
EVO R09	0.787	SOC O10	0.416	ULH O10	0.442
NUB D58	0.759	LIT OGO	0.409	ULH O05	0.441
INS 475	0.758	LIT OBO	0.401	ULA O10	0.414
PRE T13	0.63	LIT OOU	0.393	UPV O10	0.396
HUM R90	0.609	SOL L05	0.393	UPD O10	0.396
INS 525	0.515	SOL L10	0.393	ULS O05	0.361
GEA N90	0.513	LIT ODO	0.354	UMC O10	0.348
INS 625	0.506	ALT O10	0.351	UMZ O10	0.331
GEA N30	0.481	LIT OMO	0.343	UPF O05	0.31
HUM R50	0.479	ALT O03	0.339	ULS O10	0.305

3.4.2 CARTA ECOLÓGICA DE PORTUGAL (ALBUQUERQUE, 1945)

Com base no axioma "*a racionalização da Agricultura exige conhecer previamente o Meio Natural*", e por incumbência da então designada Direcção Geral dos Serviços Agrícolas, J. Pina Manique e Albuquerque, engenheiro agrónomo e silvicultor da Estação Agronómica Nacional, preparou e apresentou nas Jornadas Agronómicas de 1942 em Lisboa, a Carta Ecológica de Portugal. Em tal mapa (Anexo 4 o), o seu autor pretendeu plasmar em pormenor, uma zonagem ecológica em função do clima, a qual seria de certa forma "confirmada" pela designada *silva climática* (Albuquerque, 1945).

Com uma visão bastante concreta sobre o papel dos diversos macrofactores na definição dos principais complexos ecológicos do território nacional, Albuquerque (1945) refere três processos de fazer uma zonagem ecológica: o estudo directo dos solos, o estudo directo dos climas e, indirectamente, o estudo da vegetação indicadora. A definição de grandes tipos de paisagem climática funcionaria como a sua 1ª aproximação, num processo sintético, de âmbito geográfico. O critério pedológico corresponderia a uma zonagem feita com um segundo grau de aproximação onde, de uma forma analítica, se definiriam "micro-habitats" dentro das respectivas paisagens climáticas. Por fim, a fitossociologia, com a sua base sistemática taxonómica, serviria para a definição de "estações", designando tal termo as unidades ecológicas fundamentais (Albuquerque, 1945).

No estudo de Albuquerque (1945), foi realizada a "1ª aproximação climática", intervindo a fito-climatologia (ou fito-geografia) como adjuvante, a fim de lhe servir de contra-prova; define-se então uma *silva climática* como complexo de espécies florestais dominantes e sub-dominantes,

pelas quais se podem inferir os vários tipos de climas. Até então, a fito-climatologia portuguesa tinha sido tratada por vários autores, como os casos de Gomes (1878), Willkomm (1896), Daveau (1902), Teles (1929), Gaussen (1931) e Lautensach (1932) citados por Albuquerque (1945).

Para a zonagem ecológica de Portugal, o autor contou com 4 ordens de influências, as quais qualifica como "primaciais": atlânticas, oro-atlânticas, mediterrâneas e ibéricas. As primeiras - *atlânticas* - são dominantes sobre o Noroeste do território nacional, caracterizando-se por possuírem os maiores valores de precipitação e os mínimos do índice de aridez de Mortone $\left(\frac{100T}{P}\right)$. É o domínio da *Quercus robur* (carvalho alvarinho, pedunculado). O "epicentro" da influência atlântica situar-se-ia na Escócia e SO da Noruega (Albuquerque, 1954).

As influências *mediterrâneas* notam-se pelos valores baixos de pluviometria, em simultâneo com uma aridez excessiva, apresentando a *Prunus dulcis* (amendoeira) como característica da sua *silva climática*. Para além do "epicentro" natural do Este algarvio, o vale do Douro, na sua zona mais interior, seria também um importante polo de mediterraneidade.

As influências *oro-atlânticas* estão associadas às grandes altitudes, com os maiores valores de precipitação, resultado duma maior exposição aos ventos húmidos atlânticos; a sua representação no território português alcançar-se-ia, segundo Albuquerque (1945) no "... andar pastoril dos relevos galaico-duriense e erminiano, onde a bétula (*Betula celtiberica*) é a espécie mais característica ...".

As influências *ibéricas* far-se-iam sentir no território nacional, sem contudo e tal como as atlânticas, aí situarem o seu "epicentro". Tais influências são resultado da proximidade da meseta castelhana, duma "continentalidade excessiva", "... expressa no condomínio da azinheira (*Quercus ilex*) e carvalho negral (*Quercus pyrenaica*), e na quasi ausência da silva climática mediterrânica." (Albuquerque, 1945).

O território da província de Trás-os-Montes e Alto Douro tem a particularidade de, pela sua grande diversidade ecológica, abranger 19, ou seja, a maior parte das zonas fitoclimáticas descritas por Albuquerque (1954) para o território nacional. Nesta obra, o autor dá um destaque especial aos gradientes *atlântico-mediterrânicos* do vale do Douro, aos andares *oro-atlânticos* dos relevos mais elevados e à proximidade da *continentalidade* da meseta castelhana. As zonas abrangidas por este trabalho podem ser diferenciadas pela presença ou ausência de influências *atlânticas*. Em relação às primeiras, as influências *atlânticas* podem ser exclusivas:

- Zona atlântica sub-atlântica (AxSA)

- Zona sub-atlântica (SA)

cruzadas ou exclusivamente *oro-atlânticas*:

- Zona oro-atlântica (OA)
- Zona atlântica sub-atlântica oro-atlântica (AxSAxOA)
- Zona sub-atlântica oro-atlântica (SAxOA)

cruzadas com *ibéricas*:

- Zona ibérica sub-atlântica (IxSA)

cruzadas com *mediterrâneas*:

- Zona sub-atlântica atlântica mediterrâneo-atlântica (SAxAxMA)
- Zona sub-atlântica mediterrâneo-atlântica (SAxMA)
- Zona sub-atlântica mediterrâneo-atlântica atlante-mediterrânea (SAxMAxAM)
- Zona sub-atlântica atlante-mediterrânea (SAxAM)
- Zona sub-atlântica sub-mediterrânea (SAxSM)
- Zona mediterrâneo-atlântica atlante-mediterrânea (MAxAM)
- Zona atlante-mediterrânea (AM)
- Zona atlante-mediterrânea sub-mediterrânea (AMxSM)

e cruzadas simultaneamente com *mediterrâneas* e *ibéricas*:

- Zona sub-atlântica sub-mediterrânea ibero-mediterrânea (SAxSMxIM)

Em relação às segundas, as influências *não atlânticas* podem ser exclusivamente mediterrâneas:

- Zona mediterrânea sub-mediterrânea (MxSM)
- Zona sub-mediterrânea (SM)

ou cruzadas entre *mediterrâneas* e *ibéricas*:

- Zona sub-mediterrânea ibero-mediterrânea (SMxIM)
- Zona mediterrânea ibero-mediterrânea sub-mediterrânea (MxIMxSM)

Para a confrontação dos resultados obtidos pelo "TWINSPAN" com a Carta Ecológica, procedeu-se de forma análoga à realizada com as Grandes Regiões (Agroconsultores e COBA, 1990). O Quadro III.19, de contingência relativa às duas classificações, permite adivinhar os grandes traços da semelhança entre elas. De facto, a grande diferença no número de classes leva a que cada grupo de locais criados pelo "TWINSPAN" se manifeste em variados grupos da Carta Ecológica.

QUADRO III.19 - Tabela de contingência entre os grupos de locais definidos pelo "TWINSpan" e as classes da classificação da Carta Ecológica.

Grupos	000	001	010	011	100	101	110	111	Total
OA	0	0	0	0	0	0	3	0	3
AxSAxOA	0	0	0	0	0	0	21	0	21
SAxOA	0	0	0	0	4	0	7	0	11
AxSA	0	0	1	0	0	0	2	0	3
SA	0	0	19	11	6	5	2	1	44
IxSA	1	0	1	26	2	0	0	0	30
SAxMA	0	0	0	0	0	0	0	3	3
SAxMA	0	3	0	0	1	0	0	2	6
SAxMAxAM	0	9	4	0	9	5	0	2	29
SAxAM	1	4	11	3	0	3	0	0	22
SAxSM	6	4	12	7	0	0	0	0	29
SAxSMxIM	12	1	2	34	1	0	0	0	50
MAxAM	0	4	1	0	0	0	0	1	6
AM	4	6	0	0	0	0	0	0	10
AMxSM	5	2	0	0	0	0	0	0	7
SM	7	4	0	0	0	0	0	0	11
SMxIM	4	0	0	0	0	0	0	0	4
MxIMxSM	2	0	0	0	0	0	0	0	2
MxSM	7	2	0	0	0	0	0	0	9
Total	49	39	51	81	23	13	35	9	300

A análise canónica da tabela de contingência revela um coeficiente de correlação bastante elevado para o primeiro eixo e mediano para os seguintes (Quadro III.20). Mais uma vez, o valor de χ^2 revela a coincidência entre os vários grupos das duas classificações, apresentando um valor altamente significativo.

QUADRO III.20 - Teste de χ^2 e coeficientes de correlação canónica entre grupos de locais definidos pelo "Twinspan" e a classificação da Carta Ecológica.

	Coefficiente de correlação canónica	Valor de χ^2 associado a cada eixo	Valor de χ^2 acumulado
Primeiro eixo	0.923	256	33%
Segundo eixo	0.744	166	55%
Terceiro eixo	0.699	147	74%
Quarto eixo	0.597	107	88%
Quinto eixo	0.421	53	95%
Sexto eixo	0.340	35	99%
Total		764 (126 g.l.)	

A projecção dos grupos das duas classificações sobre os dois primeiros eixos canónicos revela que o primeiro eixo discrimina as influências oro-atlânticas das restantes, enquanto o segundo discrimina as influências mediterrânicas em relação às zonas mais atlânticas e às zonas de maior influência ibérica (figura 3.4). Confirmam-se assim as tendências oro-atlânticas do grupo 110, atlânticas do grupo 111 e ibéricas do grupo 000, distribuindo-se as mediterrâneas pelas outros grupos formados.

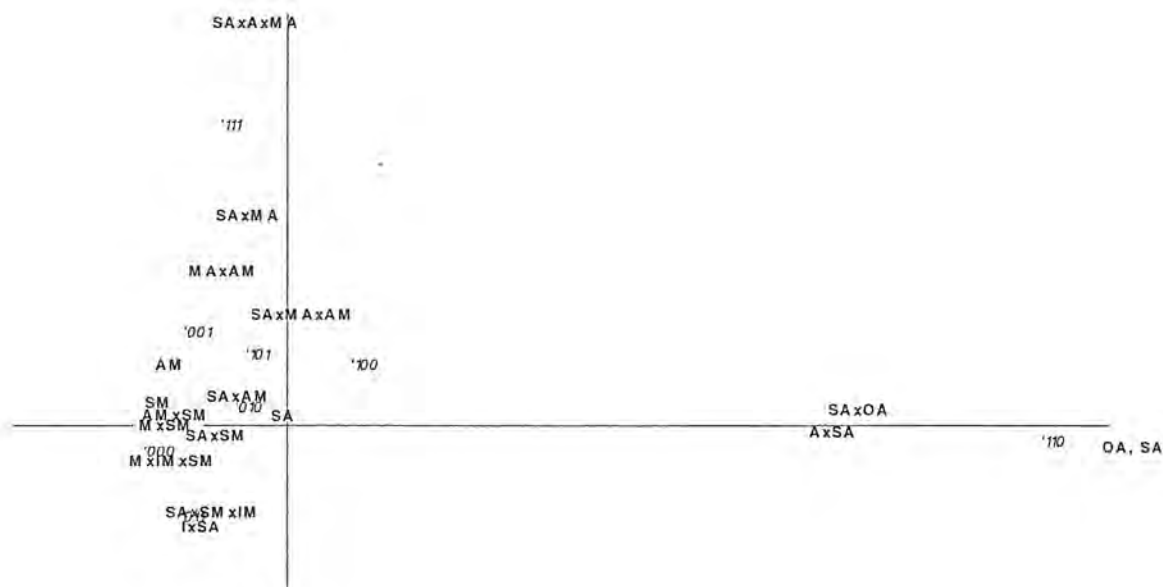


FIGURA 3.4 - Representação gráfica dos grupos de locais definidos pelo "TWINSPAN" e as classes da classificação da Carta Ecológica (Albuquerque, 1945), no espaço definido pelos dois primeiros eixos canônicos.

Tal como na classificação anterior, procedeu-se à avaliação da importância das variáveis aqui estudadas na indicação dos sectores criados pela classificação da Carta Ecológica. Foram calculados os valores de *C* e os seus resultados são os apresentados no Quadro III.21. A sua observação permite verificar mais uma vez, a importância dos factores climáticos em relação aos restantes, embora se note maior equilíbrio que nas classificações comentadas antes. Os factores fisiográficos apresentam em geral valores superiores aos de uso do solo. Tal ordem de importância confirma o descrito pelo autor em relação às aproximações efectuadas na sua elaboração da classificação; uma aproximação inicial baseada no clima, a qual é complementada com uma segunda aproximação edáfica. Nos factores fisiográficos abundam as variáveis altitudinais, o qual se deve à base altitudinal de que o autor segue para realizar a sua sectorização.

QUADRO III.21 - Valores de "*C*" mais elevados para os atributos climáticos, fisiográficos e de uso do solo, na sectorização realizada pela Carta Ecológica.

Variáveis climáticas		Variáveis fisiográficas		Variáveis de uso do solo	
EVO R09	0.689	ALT 002	0.557	ULA 010	0.487
PRE T13	0.67	SOC 030	0.553	ULH 005	0.463
NUB D58	0.645	ALT 011	0.54	ULH 010	0.45
INS 475	0.643	ALT 010	0.494	UPD 010	0.418
INS 625	0.608	ALT 001	0.493	UMZ 010	0.399
GEA N90	0.596	SOC 010	0.418	UPF 005	0.332
TMP A13	0.547	ALT 003	0.417	ULS 010	0.323
NUB D43	0.544	ALT 009	0.404	UMZ 005	0.301
TMP E09	0.539	ALT 007	0.378	UCR 005	0.297
HUM R90	0.527	RIO ABE	0.378	UMP 010	0.293

3.4.3 CARTA DE AGROTIPOS DE PORTUGAL (ALBUQUERQUE, 1945)

Na Carta de Agrotipos de Portugal (Albuquerque, 1945) são individualizados os tipos de paisagens agrárias do território nacional. Contemporânea e do mesmo autor da Carta Ecológica, foi seu objectivo principal "... individualizar as nossas lídimas «regiões» naturais, por caracteres gerais privativos, permitindo unificá-las e contrapô-las às «regiões» vizinhas;", verificando a existência ou não de espécies florestais autóctones que as individualizem (Albuquerque, 1945). Tem base fisionómica o que, tal como o seu autor refere, implica âmbito geográfico (Anexo 4 p).

Tanto a escala como o âmbito deste trabalho levou a contrapor os resultados do "TWINSPAN" com os Agrotipos definidos por Albuquerque (1945). Cada um destes tipos representa um complexo de utilizações do solo resultante dos diversos factores ecológicos que nele concorrem. Estes factores, desde os mais primários até aos mais dependentes e interligados são, no fundo, o verdadeiro objecto deste trabalho.

Trás-os-Montes e Alto Douro abrange, segundo Albuquerque (1945), quatro regiões naturais. São elas o Noroeste Cismontano, o Alto Portugal, a Beira Douro e o Nordeste Transmontano. No entanto, são as regiões do Alto Portugal e Nordeste Transmontano que ocupam a maior parte da sua área. Neles se sucedem 10 Agrotipos ou "quadros rurais", como lhes chamou também o seu autor. É da forma recorrente como eles se sucedem no território que nasce a individualização de cada uma das «Regiões» consideradas.

A região do Alto Portugal integra 3 grandes grupos de Agrotipos: o Douro, a Ribeira e a Montanha. O primeiro, o *Douro Vinhateiro*, corresponde grosso modo à conhecida zona vitícola, caracterizada pela sua monocultura instalada num tipo de fisiografia muito particular. A paisagem do tipo *Ribeira Sub-atlântica* é dividida entre as zonas onde o regadio é *dominante* e onde ele é *dominado*; são áreas de policultura, onde para além das culturas de regadio (milho, forragem), se cultivam arvenses de sequeiro, vinha e oliveira. Entre as zonas de *montanha*, Albuquerque (1945) distingue as de ocupação predominantemente *florestal* e as áreas de matos, de utilização *pastoril*. Refere também a sua base litológica em granito e xisto. Dentro deste grupo de agrotipos, inclui-se a designada *gândara*, que corresponde às zonas de encosta de influência atlântica, onde a instalação de aglomerados populacionais levou o Homem a instalar socalcos a que deu uma utilização de autoconsumo e alimentação animal.

A região do Nordeste Transmontano corresponde ao interior da província, abrigado do Atlântico pelas montanhas do Alto Portugal e Beira-Douro, e constituído por planaltos e fossas tectónicas pertencentes já à Meseta Ibérica (Albuquerque, 1945). Os agrotipos mais frequentes são designados por *Terra Quente*, *Terra Fria* e *Meia Encosta*. Nas zonas de ecologia favorável,

considera Albuquerque (1945) que predomina a *Policultura Submediterrânea*. Por *Terra Quente* entende-se a paisagem campaniça, de sequeiro extreme utilizada para seara, pousio e oliveira. Por *Terra Fria* entende-se as zonas do andar altimontano a Este do Rio Sabor, zona predominantemente dominada pela floresta de carvalho negral (*Quercus pyrenaica*), entrecortada por pequenos prados de regadio (lameiros) ou áreas de matos resultantes da degradação da floresta original. As áreas designadas de *Meia Encosta* correspondem a áreas geralmente ocupadas por sobreirais (*Quercus suber*) ou áreas resultantes da sua degradação (*cistus spp.*, etc.). As zonas de *Policultura Submediterrânea* correspondem a regiões de baixa (depressões) com predomínio do sequeiro arvense, vitícola e olivícola mas onde coexistem com o regadio de pomóideas e hortícolas.

QUADRO III.22 - Tabela de contingência entre os grupos de locais definidos pelo "TWINSPAN" e as classes da classificação da Carta de Agrotipos.

Grupos	000	001	010	011	100	101	110	111	Total
Monte Florestal	0	0	0	0	2	0	30	0	32
Monte Pastoril	0	1	19	10	12	9	4	1	56
Gândaras	0	3	0	0	0	0	0	8	11
Ribeira Verde	0	10	5	0	7	2	0	0	24
Douro Vinhateiro	0	10	0	0	0	0	0	0	10
Ribeira Perene	1	4	14	5	0	2	0	0	26
Policultura SM	4	8	9	0	0	0	0	0	21
Terra Quente	31	2	0	0	0	0	0	0	33
Meia Encosta	13	1	2	37	1	0	0	0	54
Terra Fria	0	0	2	29	1	0	1	0	33
Total	49	39	51	81	23	13	35	9	300

Tanto a região do Noroeste Cismontano como a da Beira-Douro são regiões que participam na área de estudo com alguns dos Agrotipos agora descritos para as regiões do Alto Portugal e Nordeste Transmontano.

O Quadro III.22 corresponde às coincidências dos locais da classificação feita pelo "TWINSPAN" com os Agrotipos localizados na área de estudo. Dela se adivinham as semelhanças mais marcantes entre as duas classificações. O Quadro III.23 valoriza essa semelhança, altamente significativa a avaliar pelo valor de χ^2 . Significativo é também, a correlação canónica dos dois primeiros eixos, os quais são utilizados para visualizar a localização relativa dos grupos das duas classificações (figura 3.5).

QUADRO III.23 - Teste de χ^2 e coeficientes de correlação canónica entre grupos de locais definidos pelo "TWINSPAN" e a classificação da Carta de Agrotipos.

	Coefficiente de correlação canónica	Valor de χ^2 associado a cada eixo	Valor de χ^2 acumulado
Primeiro eixo	0.900	243	27%
Segundo eixo	0.839	211	50%
Terceiro eixo	0.764	175	70%
Quarto eixo	0.717	154	87%
Quinto eixo	0.518	80	96%
Sexto eixo	0.328	32	100%
Total		895 (63 g.l.)	

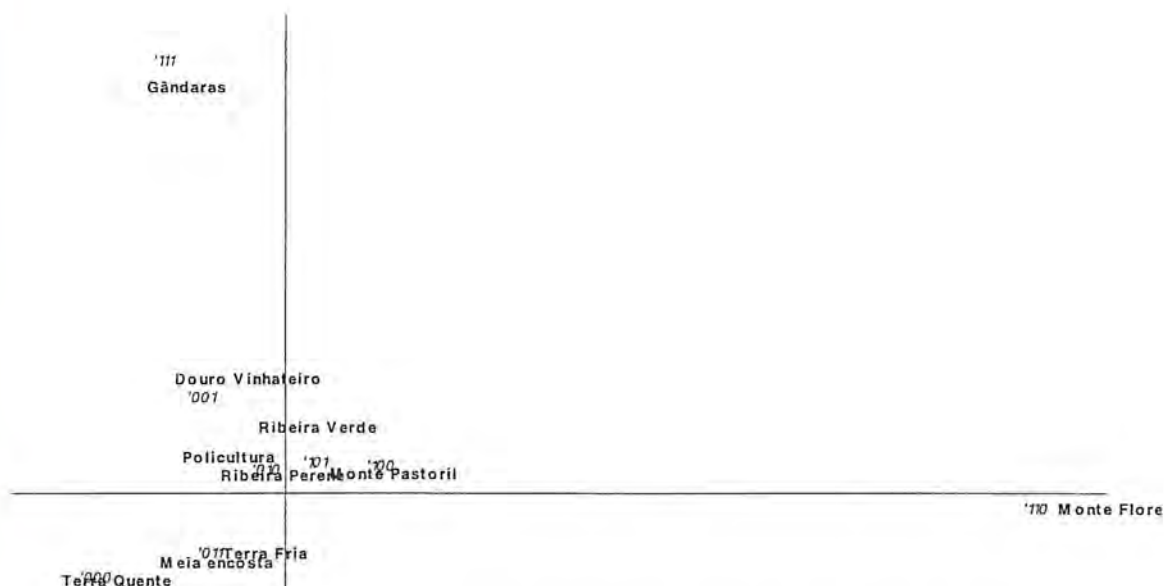


FIGURA 3.5 - Representação gráfica dos grupos de locais definidos pelo "TWINSPAN" e as classes da classificação da Carta de Agrotipos (Albuquerque, 1945), no espaço definido pelos dois primeiros eixos canónicos.

Da sua análise ressalta a discriminação entre os dois principais tipos de paisagem atlântica, tal como observado nas outras classificações com que se comparam os resultados obtidos pelo "TWINSPAN". Os restantes eixos servem para discriminar as classes relativas ao interior da província. Mais do que em outra qualquer classificação, as classes de Agrotipos de Albuquerque (1945) apresentam uma colagem quase perfeita às classes definidas pelo "TWINSPAN". A este facto não será estranho a integração de factores que empiricamente o seu autor realizou, complementando factores climáticos com factores fisiográficos e de uso ou ocupação do solo, procurando "... um mapa económico-agrário dos tipos de lavoura presente no território", que serve de complemento ao anterior (Carta Ecológica; Albuquerque, 1945).

Tal como nas classificações anteriores, foi utilizado o valor de C para avaliar o valor indicativo de cada uma das variáveis na Carta de Agrotipos. Tal como seria de esperar, os seus valores seguem os obtidos para a classificação anterior. Para tal concorrerá a íntima ligação existente entre essas duas classificações (Carta Ecológica e Carta de Agrotipos). O valor indicativo das variáveis climáticas sobrepõe-se ao das variáveis fisiográficas, e o destas ao das de uso do solo (Quadro III.24).

QUADRO III.24 - Valores de "C" mais elevados para os atributos climáticos, fisiográficos e de uso do solo, na sectorização realizada pela Carta de Agrotipos.

Variáveis climáticas		Variáveis fisiográficas		Variáveis de uso do solo	
PRE T13	0.631	ALT O11	0.530	ULA O10	0.407
NUB D58	0.589	ALT O10	0.467	UPF O05	0.377
NUB D43	0.55	ALT O01	0.463	ULH O10	0.377
INS 625	0.549	SOL L10	0.426	UMZ O10	0.327
TMP E09	0.546	ALT O03	0.400	ULH O05	0.306
GEA N90	0.545	ALT O02	0.392	UPV O10	0.287
EVO R09	0.534	ALT O09	0.371	ULS O10	0.286
INS 475	0.522	SOC O30	0.347	UPD O10	0.255
EVO R19	0.515	ALT O07	0.344	UCR O05	0.225
HUM R90	0.505	ALT O05	0.332	UMZ O05	0.223

CAPÍTULO 4

CONCLUSÕES

4.1 CONSIDERAÇÕES FINAIS

As técnicas de análise multivariada e os conceitos da teoria da informação permitiram a estruturação de um espaço conceptual onde evoluem e se plasmam os diferentes componentes, as inter-relações e respectivo funcionamento ecológico do sistema territorial transmontano.

De entre as técnicas de análise multivariada que foram abordadas e confrontadas, as técnicas de classificação permitem a individualização dos principais componentes do sistema ecológico inerente ao território da província transmontana; a natureza deste objecto de estudo implica a utilização de uma classificação divisiva, para distinguir as principais "fronteiras" desse sistema ecológico. O método "TWINSPAN" tem dado provas da sua excelência na "compreensão" das principais tendências dos sistemas naturais que analisa, adequando-se às características que são inerentes a esses sistemas.

As técnicas de análise de concentração utilizadas avaliam a convergência ou divergência existente entre esses componentes, assim como as coincidências do sistema agora definido, com anteriores abordagens do tema. Estão neste caso a análise da coincidência entre os componentes agora definidos e os idealizados em zonagens realizadas anteriormente, a sua quantificação e evolução em relação às técnicas de recolha de dados que se utilizou.

Os conceitos da teoria da informação, e seus desenvolvimentos posteriores aplicados às ciências naturais em geral, e à ecologia em particular, permitem a interpretação do papel de cada componente desse sistema conceptual. É assim possível indicar os grandes protagonistas do funcionamento ecológico do território, cujo papel é fundamental na individualização das diferentes ecologias e na sua organização ao longo do território.

A organização funcional do território transmontano adapta-se à sua conformação física, ao qual o Homem foi proporcionando modelos de utilização do espaço específicos a cada situação particular. Fica demonstrado que a variação Litoral ↔ Interior (Oeste - Este), traduzida fundamentalmente em diferentes situações climáticas, é o factor mais importante na discriminação ecológica do território. À influência oceânica sobrepõe-se a história geológica, ajustando-se as zonas que compõem o sistema montanhoso galaico-duriense às de maior influência atlântica, e as

predominantemente formadas pelo alongamento transmontano da meseta centro-ibérica às de menor influência atlântica, com clima do tipo mediterrânico.

A variação altitudinal é o segundo factor mais importante na classificação do território em estudo. Na região de influência atlântica distingue as regiões de montanha das dos vales mais importantes. Na região interior distingue os vales do Douro e seus afluentes, da região planáltica correspondente à meseta residual.

Para além da distinção litoral/interior e altitudinal, distinguem-se e individualizam-se oito tipos de paisagens que formam a matriz da organização do espaço transmontano. As regiões de baixa altitude do interior da província estão associadas aos vales dos rios que a sulcam. O desenvolvimento e orientação dominante dessas zonas húmidas, dispostas longitudinalmente, induzem condições climáticas e físicas diversas e, naturalmente, uma diferente organização do espaço. No *vale do Douro*, de fisiografia abrupta mas relativamente suave em termos de clima, com um mesoclima protegido e propício à penetração da influência mediterrânea, o homem ergueu socalcos que levaram a que a partir de solos esqueléticos se "fabricasse" um substrato pedológico que permitiu a instalação da cultura da vinha com sucesso. Os restantes *vales interiores*, mais largos e aplanados, mas mais secos, são aproveitados predominantemente para o cultivo de cereais de sequeiro.

As regiões mais altas do interior da província distinguem-se fundamentalmente pela sua base litológica. O *interior granítico* surge associado à periferia do vale do Douro e seus principais afluentes, apoiado em relevos residuais do maciço ibérico ou áreas marginais do sistema montanhoso galaico-duriense; de clima idêntico ao do vale do Douro, apresenta litologia granítica patente em frequentes afloramentos rochosos, entre os quais se desenvolveram cambissolos aproveitados para vinha e leptossolos ocupados maioritariamente por povoamentos florestais à base de pinheiro bravo. As restantes áreas da região interior constituem o *planalto cerealífero*, o maior e mais fundamental espaço da província; de litologia xistosa ou básica, e declives relativamente suaves, o solo é aproveitado para o cultivo de cereais de sequeiro, aproveitando-se as depressões mais favoráveis em termos de humidade para a produção de pastagem e/ou forragem, igualmente de sequeiro.

Nas regiões de mais baixa altitude do sector mais "atlântico" da província, é possível separar as zonas de meia encosta das localizadas no fundo dos vales dos rios da região. As primeiras apresentam um relevo com carácter de *montanha média*, de litologia xistosa e sem uma ocupação dominante do solo, à qual apenas é possível associar algum cultivo de cereal. Por outro lado, o fundo dos vales apresenta-se estratificado em socalcos largos onde se desenvolveram solos

antrópicos que são a base de culturas de regadio. Pequenos bosquetes de pinheiro bravo nos locais mais desfavoráveis completam este padrão paisagístico de *ribeira*.

As regiões de montanha da região atlântica da província dividem-se entre as *cumeadas graníticas* e aplanadas, e as *gândaras* mais protegidas que o Homem escolheu para instalar a sua habitação. As primeiras, de clima rigoroso e ocupadas predominantemente por vegetação arbustiva, servem de base a uma actividade pastoril já ancestral. As segundas apresentam o espaço aproveitado por socalcos de dimensão média circundando pequenos núcleos urbanos, e são aproveitados para a prática de uma agricultura de auto-sustento, de apoio às famílias que vivem da pastorícia praticada nas *cumeadas* antes descritas.

A subsectorização destas regiões, se bem que possível, necessita de um tratamento individualizado, com estudo de factores próprios, inerentes a uma maior escala de trabalho, com o correspondente nível de precisão e rigor. Tal abordagem serviria para o estudo aprofundado dos modelos de organização territorial, sua evolução, momento presente e perspectivas futuras.

A informação disponível sobre a ecologia da área em estudo e a experiência de metodologias de inventário ambiental, tanto própria como publicada anteriormente, permitiu a idealização de um sistema de recolha de dados simples e facilmente transponível a outras regiões. Um sistema de transectos rectilíneos, apoiados sobre os meridianos que atravessam a província, permitiu localizar com facilidade pontos de amostragem, que de uma forma sistemática representam a integridade ambiental do território.

O esquema de amostragem sobre meridianos, a chave de todo o processo de aquisição de informação, foi sempre o mesmo, não se confrontando com outras possíveis variações de intensidade e localização dos locais de inventário. Os resultados obtidos seguem a unanimidade dos autores que se debruçaram sobre o tema, levando a que não se optasse pela sua modificação. As classes obtidas, embora em número bastante diverso das anteriores classificações do território, coincidem com as mais importantes classes definidas por outros autores.

Entre as classes agora definidas e as anteriormente citadas por outros autores, ocorrem dois tipos de situações; na maior parte das situações, várias classes de uma classificação anterior confluem em apenas uma classe da classificação do "TWINSPAN", ou em menor proporção, nenhuma destas classes representam o ambiente que caracteriza alguma classe definida pelas anteriores classificações.

Na classificação das Grandes Regiões, ambientes semelhantes com diferentes localizações são classificados de forma diversa; tal acontece com as grandes regiões Tua e Alto Douro que confluem na classe 001 (*Vale do Douro*), e com Bragança e Miranda-Mogadouro que confluem na classe 011

(*Planalto Cerealífero*). Também na Carta de Agrotipos, a Meia Encosta e Terra Fria representam a quase totalidade da classe 011 (*Planalto Cerealífero*). Tal situação é devida sobretudo à especificidade de cada classificação, assim como ao nível de detalhe usado na sua elaboração, características não compatíveis com as escalas deste trabalho (1/500 000 e 1/100 000). Na Carta Ecológica, o seu grande número de zonas leva a que algumas delas, pela sua fraca representatividade, não se identifiquem com as classes criadas pelo "TWINSPAN"; estão neste caso as zonas ecológicas OA, A x SA, MA x AM, AM x SM, SM x IM e M x IM x SM.

Das aproximações realizadas no passado ressalta a importância das influências atlânticas e altitudinais na discriminação ecológica da província. Tal é o fundamento da primeira carta publicada por Albuquerque (1945) - Carta Fito-edafo-climática ou Carta Ecológica - realizada com base em factores biofísicos, (clima e altitude, fundamentalmente). A Carta de Agrotipos, desenvolvida com base na anterior, consubstancia a importância dos diversos modelos de organização do espaço, traduzindo-se numa carta de alto valor informativo, que até hoje tem sido injustamente esquecida. É com ela que as sectorizações definidas por este trabalho mais se identificam, apresentando uma razoável convergência de resultados. Sendo datada de 1945, seria de todo oportuno a sua actualização, tanto como estudo da evolução recente da utilização do território, como base de trabalho para a definição de zonas operacionais de políticas específicas de desenvolvimento, de modo harmonioso e convergente com o seu funcionamento natural.

4.2 PRINCIPAIS RASGOS DA ECOLOGIA DE TRÁS-OS-MONTES E ALTO DOURO

A comparação das classificações existentes para a província de Trás-os-Montes e Alto Douro com os resultados obtidos pelo "TWINSPAN", confirma a importância da influência climática Atlântica e Continental na sectorização ecológica do território transmontano. A análise dos atributos com maior responsabilidade na definição dos sectores representados em cada carta permite inferir a natureza das diferentes classificações. As maiores semelhanças entre as quatro classificações verificam-se nas variáveis climáticas e nas de uso do solo. Entre as climáticas, encontram-se os atributos caracteristicamente atlânticos, como a precipitação, a nebulosidade, a humidade relativa e o número de dias de geada elevados, e a evaporação baixa. As variáveis de uso do solo mais importantes repartem-se de forma equitativa entre as de características atlânticas e as de características continentais, os "prados permanentemente húmidos" e "de altitude" para os primeiros, e os "pomares de amendoeira" como dominante e de "fruteiras" como dominada para o segundo. As maiores divergências verificam-se ao nível dos atributos fisiográficos, onde apenas os níveis altitudinais dos 300 metros para os gradientes de clima mediterrâneo, e de 1000 metros para os atlânticos, coincidem em todas as classificações. A principal divergência acontece na

classificação das Grandes Regiões, que apresenta uma sectorização muito dependente da litologia (cinco litologias entre os 10 primeiros C's de variáveis fisiográficas). Entre as outras 3 classificações, as coincidências são maiores, nomeadamente em relação aos níveis altitudinais (atlânticos entre 900 e 1100 metros de altitude e mediterrâneos com 100 metros), e ao tipo de armação de terreno (caracteristicamente atlântico com socalcos largos e caracteristicamente mediterrâneo com socalcos apertados).

A quantificação da semelhança entre as quatro classificações fez-se com base na classificação dos locais de amostragem realizada por cada uma dessas classificações. A escolha de índices para a avaliação das diferenças existentes entre as diversas classificações em confronto deriva da própria natureza dos dados trabalhados; são dados qualitativos, do tipo presença/ausência (binários), em que a ausência simultânea de um carácter não é relevante. Para tais situações, Legendre e Legendre (1984) propõe os índices de Sørensen, Ochiai e Kulczynski, entre outros. O software disponível - SYNTAX IV (Podani, 1990) - permite avaliar a dissimilaridade entre as classificações consideradas, de acordo com os dois primeiros índices referidos, os quais são os mais frequentemente utilizados em trabalhos do género. A fórmula de Sørensen (1948; citado por Legendre e Legendre, 1984), transformada em distância pela subtração do seu valor à unidade, é:

$$1 - \frac{2a}{(2a+b+c)}$$

A fórmula de Ochiai (1957; citado por Legendre e Legendre, 1984), transformada da mesma forma que a anterior, é:

$$1 - \frac{a}{\sqrt{(a+b)(a+c)}}$$

Assim, para cada classificação dos 300 locais de amostragem é definida uma matriz de incidência 300X300, cujos coeficientes (c_{ij}) representam se o respectivo par de locais pertence (1) ou não (0) à mesma classe. Quando se comparam duas classificações, são contabilizados o número de pares de locais que coincidem na mesma classe em cada uma das classificações comparadas (a), o número de pares de locais que coincidem na mesma classe apenas na primeira classificação (b), e o número de pares de locais que coincidem na mesma classe apenas na segunda classificação (c). Os valores apresentados são "standartizados", ou seja, representam o quociente entre os valores obtidos pelas fórmulas acima e o máximo possível em função do número de classes e pares de locais em causa; esbatem-se assim as diferenças inerentes ao facto das classificações contemplarem um número de classes diferente.

Da análise dos resultados do quadro IV.1, podemos constatar que:

- os valores obtidos para os dois índices pouco diferem, o que já seria de esperar pela semelhança das suas fórmulas;
- as maiores semelhanças são as existentes entre as classificações afins, como são os casos das Cartas Ecológica e de Agrotipos;
- as maiores diferenças dizem respeito a comparações com as Grandes Regiões, ao qual não será estranho as particularidades desta classificação (zonagem, espaço contínuo).

Para além dos aspectos referidos, há a notar a coincidência entre a Carta de Agrotipos e a classificação obtida pelo "TWINSpan". A tal facto não deverá ser alheio a coindidência entre os aspectos considerados na elaboração da Carta de Agrotipos (não estritamente climáticos), e as variáveis utilizadas neste trabalho. De notar também a pequena aproximação entre a classificação feita pelo "TWINSpan" e a carta de Grandes Regiões, o qual derivará do facto de esta retratar áreas de identidade individualizada, considerando muitos aspectos que se sobrepõe aos meramente climáticos.

QUADRO IV.1 - Coeficientes de dissemelhança entre as 4 classificações em confronto (coeficientes de Sørensen e de Ochiai).

<u>Coeficiente de Sørensen:</u>				<u>Coeficiente de Ochiai:</u>			
	Carta ecológica	Carta de agrotipos	Grandes Regiões		Carta ecológica	Carta de agrotipos	Grandes Regiões
Carta de agrotipos	0.43860			Carta de agrotipos	0.43391		
Grandes Regiões	0.74660	0.75387		Grandes Regiões	0.74696	0.75040	
"TWINSpan"	0.71256	0.63120	0.68431	"TWINSpan"	0.70133	0.62615	0.66645

A semelhança entre os resultados do "TWINSpan" e as classificações da província transmontana, em cada um dos eixos de amostragem, permite conhecer a forma como o delineamento de inventário considerado recolhe a informação ambiental do território. A figura 4.1 representa graficamente o valor do índice de dissemelhança de Sørensen, obtido ao comparar as classificações existentes com a obtida pelo "TWINSpan", para cada um dos meridianos considerados. A sua observação permite constatar que a aproximação entre as classificações existentes e a realizada pelo "TWINSpan" varia consoante as classes definidas para o território em causa. No Quadro IV.2 são apresentados os valores de dissemelhança mais baixos, entre as classes das classificações existentes e as da realizada pelo "TWINSpan".

Os valores mais baixos de dissemelhança por meridiano são obtidos na longitude 7°45', a região onde há unanimidade entre os diversos autores no delineamento da fronteira entre o ambiente de clima marcadamente atlântico e marcadamente mediterrânico. O eixo montanhoso galaico-duriense e o início do Douro Vinhateiro, fazem com que se aproximem os limites das classes Monte

Florestal (Carta de Agrotipos), A x SA x OA (Carta Ecológica) e Barroso (Grandes Regiões) com a classe 110 do "TWINSPAN", a classe Gândaras (Carta de Agrotipos) com a classe 111 do "TWINSPAN", a classe Douro (Grandes Regiões) com a classe 001 do "TWINSPAN", e a classe Beira Douro (Grandes Regiões) com a classe 101 do "TWINSPAN".

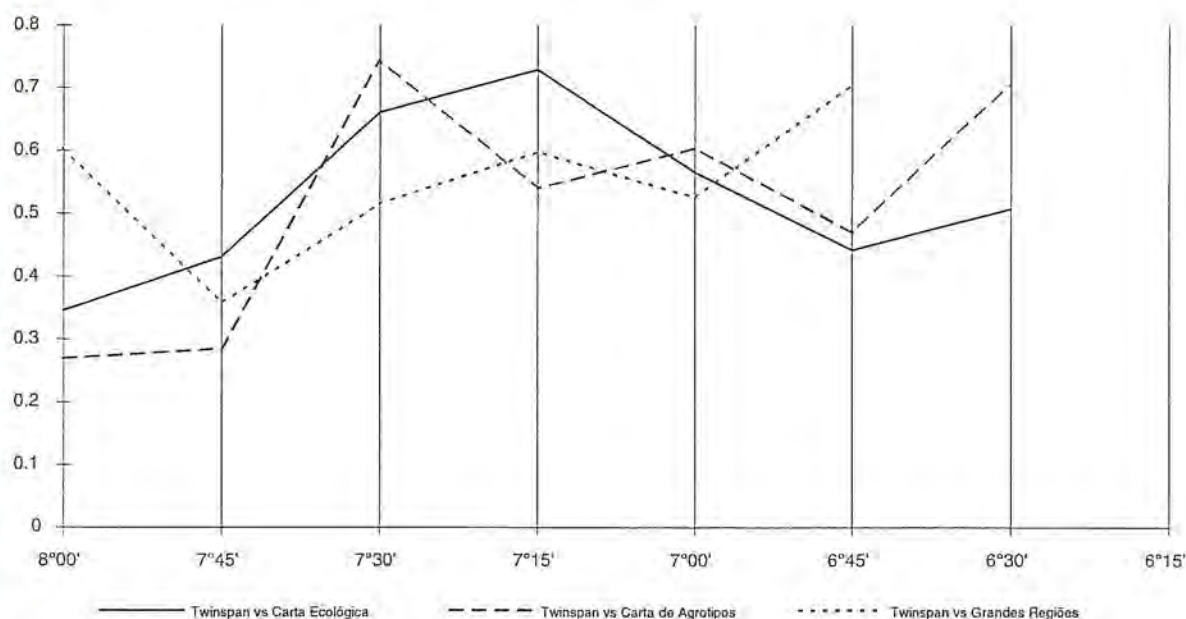


Figura 4.1- Representação gráfica dos valores do índice de Sørensen de dissimilância entre a classificação do "TWINSPAN" e outras com as quais é comparada, para cada meridiano considerado na recolha de dados.

Os meridianos da região central da Província são os que apresentam maiores divergências entre a classificação "TWINSPAN" e as outras três com as quais é comparada. A classificação agora realizada, comparada com a das Grandes Regiões, apresenta valores relativamente baixos nos eixos 7°30' e 7°00', aos quais não será estranho a ocorrência da região do Douro, bastante coincidente com a classe 001 do "TWINSPAN". A Carta de Agrotipos também se identifica relativamente bem com a classificação do "TWINSPAN", apresentando valores relativamente baixos nos eixos 7°15' e 6°45', concorrendo para tal a coincidência do Agrotipo Terra Quente com a classe 000 do "TWINSPAN", e dos Agrotipos Meia Encosta e Terra Fria com a classe 011 do "TWINSPAN". Por último, sublinha-se a semelhança entre a classificação realizada e a Carta Ecológica no eixo 6°45', a qual se atribui à sobreposição da região ecológica SA x SM x IM, com o que o "TWINSPAN" discriminou como classe 001.

Uma análise aos valores da dissimilância entre as classes agora definidas e as das classificações realizadas anteriormente, permitem inferir várias traços caracterizadores da agora alcançada. A classe 110 é a única que é claramente referenciada por todas as outras divisões, nomeadamente o Agrotipo Monte Florestal, a zona ecológica A x SA x OA e a Grande Região Barroso. A individualidade de tal classe vinha já sendo revelada pelo alto valor informativo das

variáveis, fundamentalmente climáticas que a caracterizam (Quadro III.14). Por outro lado, a identificação entre classes é mais elevada entre as "extremas" (000, 001, 110 e 111), o que denota bem a importância dada pelos diversos autores à diversidade ambiental da província, particularmente ao contraste climático entre Atlântico e Mediterrâneo. No entanto, é possível verificar a importância dada na Carta de Agrotipos e Carta Ecológica à classe 011, a maior em número de locais (81). Pelo seu elevado número de locais de amostragem, poderia pensar-se ser difícil a identificação dessa classe (011) com alguma das outras classificações em particular; tal classe, representativa da região que genuinamente caracteriza Trás-os-Montes e Alto Douro (planalto cerealífero de origem xistosa), é bem "repartida" entre os Agrotipos Meia Encosta e Terra Fria, assim como se torna preponderante a sua identificação com a zona ecológica SA x SM x IM. A classe 010 é a única que não apresenta uma dissemelhança inferior a 0.5 com qualquer das classes consideradas pelas outras classificações. O seu carácter disperso e pontual (regiões florestais sobre granitos) não deverá permitir a sua identificação com as diversas zonagens consideradas.

QUADRO IV.2 - Coeficientes de dissemelhança inferiores a 0.5 entre as classes da classificação do "TWINSPAN" e as classes das outras classificações em confronto.

CLASSES DA CLASSIFICAÇÃO DO TWINSPAN	000	001	010	011	100	101	110	111
CARTA DE AGROTIPOS								
Monte Florestal							0.104	
Gândaras								0.200
Terra Quente	0.204							
Meia encosta				0.452				
Terra Fria				0.491				
CARTA ECOLÓGICA								
A x SA x OA							0.250	
SA x SM x IM				0.481				
GRANDES REGIÕES								
Douro		0.239						
Barroso							0.333	
Beira Douro						0.333		
Tâmega					0.478			

Os valores de dissemelhança obtidos entre as classes do "TWINSPAN" e as das classificações existentes permitem inferir sobre a sensibilidade daquela classificação aos gradientes apontados pelos autores referidos. Os valores considerados (dissemelhança inferior a 0.5) ocorrem entre classes do "TWINSPAN" e classes que representam 50% dos Agrotipos (54% dos locais), 10% das Zonas Ecológicas (23% dos locais) e 31 % das Grandes Regiões (34% dos locais), que foram descritas para a Província. Tais valores vão de encontro aos valores de Sørensen para a globalidade da convergência entre as classificações.

SÍNTESE

I - REVISÃO DE METODOLOGIAS DE ANÁLISE DE DADOS

Métodos de análise multivariada

As metodologias de análise multivariada, conjuntamente com o cálculo matricial, a análise dimensional e a análise de informação, são ferramentas matemáticas utilizadas hoje em dia para a investigação, a compreensão e a modelização de sistemas ecológicos. O seu interesse advém da possibilidade de considerar os sistemas naturais como conjuntos de objectos caracterizados por um elevado número de factores (variáveis), mais ou menos interrelacionados entre si, e com tipos distintos de respostas à evolução dos ecossistemas.

O desenvolvimento das técnicas multivariadas data do princípio do século, mas foi o uso dos computadores a razão de essas técnicas, nos últimos trinta anos, se tornarem acessíveis e potentes na análise de grandes quantidades de informação. Hoje em dia diferenciam-se entre métodos de classificação e de ordenação, ambos baseados na análise de semelhança entre objectos: os primeiros servem para formar grupos de objectos com uma identidade bem definida em função das respectivas variáveis; os segundos servem para determinar a organização ou disposição relativa dos objectos, tendo em conta as coincidências das variáveis neles mesmos.

Os métodos de classificação, em função do desenvolvimento do seu algoritmo, definem-se como aglomerativos e divisivos; os primeiros aplicam-se aos casos de objectos perfeitamente definidos e individualizados, e o seu procedimento aglomerativo traduz a evolução comum desses indivíduos; os divisivos tentam definir fronteiras num universo que se apresenta como um *continuum*, determinando subdivisões com identidade bem vincada. As classificações dizem-se hierárquicas quando a aproximação ao seu resultado se faz de uma forma progressiva, de níveis inferiores a superiores (aglomerativas) ou vice-versa (divisivas). Consoante o número de variáveis tidas em conta nos processos classificativos, as técnicas de classificação dizem-se monotéticas ou politéticas.

Os procedimentos aglomerativos de classificação distinguem-se pelo que determina a fusão de objectos e/ou seus grupos; podem-se ter em conta as distâncias entre eles, a homogeneidade dos grupos formados, o aumento da informação da classificação no seu conjunto, etc. De fundamental importância é a escolha do índice que mede a semelhança entre objectos e/ou grupos, de acordo com a natureza dos dados, nominais, ordinais ou reais. Os procedimentos divisivos servem-se da

presença/ausência de determinadas variáveis para repartir o conjunto a classificar; a selecção dessas variáveis pode ser feita por critérios de informação mútua, valores médios de qui-quadrado ou ainda pela preponderância com que ocorrem em determinado extremo de um eixo de ordenação dos objectos a classificar.

Os métodos de ordenação utilizam-se em análises directas ou indirectas de gradientes dos dados em estudo. Os seus procedimentos baseiam-se na construção de uma matriz que traduz a semelhança entre os objectos, a qual é transformada por decomposição espectral. Os métodos directos pressupõem um prévio conhecimento da estrutura dos dados: a repartição prévia dos objectos por grupos identificados com algum factor ou a distinção das variáveis entre independentes e dependentes. Entre as técnicas mais utilizadas encontra-se a análise discriminante e a das correlações canónicas; a primeira indica aquelas variáveis que melhor reconhecem as diferenças entre os grupos de objectos previamente definidos; a segunda quantifica a responsabilidade da resposta de um grupo de variáveis definido como dependente, a outro considerado como independente.

Os métodos indirectos são uma aproximação à estrutura organizativa dos dados. Entre as técnicas de análise indirecta encontram-se as componentes principais, a análise factorial, a de correspondências e a de coordenadas principais; as suas diferenças residem no tipo de dados que consideram, na variação de dados que analisam, e na ordenação simultânea de objectos e variáveis. Os seus resultados indicam eixos definidos em função da participação ou peso das variáveis utilizadas na análise, a respectiva importância na explicação da variação dos dados, e a localização dos objectos no espaço multidimensional definido por esses eixos.

Teoria da Informação

Os dois conceitos antagónicos de ordem e entropia, já enunciados na segunda lei da termodinâmica, são hoje em dia a base de desenvolvimento de modernos conceitos ecológicos. A sua quantificação por Shannon (1948, citado por Legendre e Legendre, 1984b), como o simétrico da quantidade de informação de uma mensagem, criou a teoria da informação e transformou a entropia no parâmetro de dispersão por excelência de variáveis do tipo nominal, tal como as que caracterizam o meio natural. As metodologias da teoria de informação aplicam-se hoje à modelização de sistemas, estimativa da diversidade e análise estatística de complexos de dados ecológicos.

Margalef (1957, citado por Daget, 1979) conceptualizou a proporção de espécies numa amostra de determinada comunidade como uma mensagem sobre a estrutura e composição dessa

comunidade. Deste modo, o seu valor utiliza-se hoje para a quantificação de propriedades estruturais, diversidade, mutualidade e equivocação dos sistemas ecológicos. A quantidade de informação partilhada e interactiva entre duas ou mais variáveis é também uma forma de avaliar a variação conjunta de distribuições de frequência do tipo nominal, o que exprime a sociabilidade de espécies, interdependência de factores, etc. (Daget e Godron, 1982).

II - ECOLOGIA DE TRÁS-OS-MONTES E ALTO DOURO

Área de Estudo

Trás-os-Montes e Alto Douro, província do Nordeste de Portugal, resulta do prolongamento da meseta central ibérica cortada a Este e Sul pelo Vale do Douro e a Oeste pelas Montanhas do Maciço Galaico-Duriense. Ribeiro e Lautensach (1987) descreveram-na como o Norte Transmontano, uma paisagem cerealífera entremeada por afloramentos rochosos e pequenos bosques de *Quercus pyrenaica*. No século XVII começou a utilizar-se a parte do vale do Douro do Sul da Província para o cultivo da vinha em pequenos terraços, erguendo-se uma das mais espectaculares paisagens agrárias do mundo.

Objectivos

A selecção da região de Trás-os-Montes e Alto Douro para a realização deste estudo deve-se, antes de mais, à disponibilidade de informação sobre as variáveis da sua ecologia; a cartografia dos solos da província terminou mesmo antes do começo deste trabalho, recolhendo grande quantidade de elementos biofísicos e bioclimáticos da província. Pretendeu-se investigar os tipos de ambiente que ocorrem na província, assim como a capacidade de resolução de um tipo de amostragem regular em linhas paralelas, ao ser comparado com cartografias temáticas existentes. O território entende-se como um *continuum* no qual se definem fronteiras para permitir a sua compreensão e caracterização. Com uma perspectiva ecológica - *ecologia da paisagem* -, e à luz da teoria geral de sistemas, pode-se considerá-lo estruturado hierarquicamente, como um *holon* de sistemas e subsistemas; a sectorização do espaço natural implica a consideração de múltiplos factores operativos e interrelacionados, convergentes na definição dos diversos gradientes em que ocorrem.

Amostragem

Foram investigados 300 pontos de referência situados sobre 8 eixos da rede geodésica internacional. Os pontos de amostragem localizaram-se em cada minuto da latitude da província, sobre as longitudes 8°00', 7°45', 7°30', 7°15', 7°00', 6°45', 6°30' e 6°15'; os eixos, escolhidos por

facilidade da repetitividade da amostragem (referência geográfica), coincide no entanto com as maiores tendências da variação bioclimática: o Norte de influência sobretudo atlântica, e o Sul de características climáticas mediterrâneas.

A classificação ecológica que se realiza a partir da referida amostragem entende-se como definição, com critério múltiplo, de "tipos de paisagem", aos que se atribui uma potencialidade produtiva ou de uso característica. As variáveis usadas para esta classificação agrupam-se em climáticas, fisiográficas e de uso do solo.

As variáveis climáticas foram obtidas da caracterização realizada na Carta de Solos, cobrindo aspectos macro e mesoclimáticos. Estão nestes casos as temperaturas média, mínima e máxima anuais, sua amplitude, o número de dias de geada, humidade relativa, precipitação, insolação, nebulosidade, evaporação e evapotranspiração.

As variáveis fisiográficas foram seleccionadas a partir da Carta Topográfica, Carta de Solos e Carta Geomorfológica, recolhendo na medida do possível todos os aspectos que caracterizam a base física do espaço transmontano. Estão neste caso a altitude, o tipo de relevo, o declive, a litologia, os aspectos geomorfológicos de âmbito local, a importância dos afloramentos rochosos, o solo e a presença de muros de suporte de terras.

Com as variáveis do uso do solo tenta-se considerar os diversos modelos de utilização do espaço que foram sendo elaborados e melhorados ao longo dos tempos. Foram usadas as classes de "uso actual do solo" da Carta de Solos. Aí, são referidas as culturas de sequeiro intensivo e extensivo, as culturas de regadio, as culturas da oliveira, da vinha, da amendoeira, do castanheiro e outras fruteiras, os pastos de montanha e pastagens húmidas e de sequeiro, os bosques de carvalhos, de sobreiro e azinheira, pinhais e matos; considerou-se tanto a sua ocorrência como uso dominante (superior a 50% da área) ou dominado (inferior a 50%) na unidade cartográfica correspondente a cada ponto de amostragem, pretendendo-se assim incorporar no estudo aspectos complementares importantes na compreensão da dinâmica da paisagem.

Todas as variáveis consideradas foram transformadas em variáveis do tipo presença/ausência, utilizando-se as classes definidas pelos autores da sua cartografia de base (Agroconsultores e COBA, 1990; Ferreira, 1981); esta foi a forma de considerar ao mesmo tempo atributos do tipo quantitativo real (temperatura, altitude, etc.), atributos ordinais (uso do solo dominante ou dominado) e atributos do tipo qualitativo nominal (litologia, solo, etc.). Foram assim consideradas 54 variáveis climáticas, 39 fisiográficas e 28 de uso do solo.

Análise da Informação

Com a metodologia de análise pretende-se, além de classificar os pontos de amostragem, valorizar a capacidade descritiva deste tipo de amostragem efectuado em termos das sectorizações prévias. Também é possível avaliar a informação transmitida por cada uma das variáveis aos sectores ou grupos definidos. Realizaram-se análises de classificação hierárquica divisiva, de concentração por correlações canónicas e pela teoria da informação.

A classificação fez-se pelo método TWINSPAN (Hill, 1979), uma técnica hierárquica divisiva e politética, desenvolvida para classificação de inventários de vegetação, hoje em dia com múltiplas aplicações em ecologia e outras ciências. O seu emprego justifica-se pela própria natureza do objecto de estudo: o território. A classificação pelo TWINSPAN baseia-se na divisão do primeiro eixo de ordenação canónica nos objectos em estudo; dependendo da frequência com que estes aparecem nas duas metades do plano factorial, identificam-se atributos definidores (exclusivos de um dos extremos) e preferenciais (3 vezes mais frequentes de um dos lados); tal definição permite a reordenação dos objectos, em função do estado definidor e/ou preferencial dos atributos que caracterizam cada um deles, sendo essa nova ordenação a que definitivamente define a partição dos objectos. Cada conjunto, formado pela divisão do conjunto inicial, é por sua vez dividido noutros dois, de forma análoga à descrita, um processo que se repete até ao nível definido pelo investigador. Os grupos de objectos formados permitem, em função dos seus atributos, o agrupamento destes. Os resultados apresentam-se numa tabela de dupla entrada de grupos de objectos por grupos de atributos.

Vários autores comentam as vantagens e inconvenientes da utilização do "TWINSPAN" (Gauch, 1982; Jongman e col, 1987; Vos e Stortelder (1992)); entre as vantagens referem a possibilidade de utilizar dados originais, a classificação integrada de objectos e dos seus atributos, a ordenação dos grupos (Gauch, 1982); entre as segundas citam a linearidade das suas respostas (que não costuma ocorrer na natureza), a sobrevalorização de objectos pouco informativos, e a subestimação dos dados intermédios.

Resultados

Utilizando três níveis de partição das localidades do estudo, foram definidos 8 grupos representando outros 8 tipos de paisagem de Trás-os-Montes e Alto Douro; os atributos definidores e/ou preferenciais que lhes correspondem ajudam a entender a sua identidade. O primeiro nível divisivo separa os pontos caracterizados pela influência climática atlântica dos de influência climática mediterrânea, discriminados por factores exclusivamente climáticos: amplitude térmica,

nebulosidade, etc. De entre as variáveis preferenciais, para além de outras climáticas que confirmam as definidoras anteriores, aparecem a "altitude média de 700 metros" e as "culturas de cereais de sequeiro extensivo" e "olivais" para os sectores mediterrâneos, e "declives", "afloramentos rochosos", "cambissolos pouco abundantes" e "pinhais" para os sectores atlânticos. A análise das variáveis permite inferir a separação entre o planalto cerealífero transmontano das montanhas e declives atlânticos do eixo galaico-duriense.

O segundo nível separa, no grupo mediterrânico, os vales do Douro e dos seus afluentes principais, das zonas planálticas mais características da província; os seus factores definidores e preferenciais são a temperatura, evaporação, evapotranspiração e humidade, dias de geada, altitude e declive, assim como as culturas da vinha e da oliveira. A uma paisagem de clima ameno, mas seco, de declives acentuados nas quais se cultiva a vinha e a oliveira, contrapõe-se o planalto de climatologia mais extrema, maiores precipitações e menores evapotranspirações, de relevo ondulado e predominantemente ocupado por cambissolos. No grupo atlântico separam-se as zonas de montanha das de vale mais ou menos estreito. São factores de temperatura e precipitação, evapotranspiração e humidade, altitude e matos, o que descreve a identidade destes novos sectores. As zonas de vale ("ribeira") tem clima mais suave, a montanha é mais complexa, abrupta e selvagem (com abundância de matos).

O terceiro nível divisivo ressalta as características dos grupos anteriormente definidos, individualizando sectores em cada um deles, mediante um processo em que as variáveis de fisiografia e de solo vão ganhando importância informativa. No grupo dos vales mediterrânicos separa-se o do Douro dos dos seus afluentes; no Douro, o clima é mais suave e a vinha é claramente dominante, localizada em socacos erguidos para permitir a ocupação dos seus declives abruptos; nos vales dos afluentes, de relevo ondulado e mais áridos que o anterior, cultiva-se preferencialmente cereais de sequeiro extensivo. Do grupo do planalto separam-se sectores que se diferenciam pela sua litologia e topografia, um grupo granítico e rochoso, utilizado para vinha e pinhal, e outro que agrupa xistos e rochas básicas, nos que se intercalam zonas secas de cereais e zonas mais húmidas com pastagens de sequeiro. Na parte atlântica, no grupo das "ribeiras", separam-se as suas zonas mais estreitas e húmidas, organizadas em socacos largos com culturas de regadio, das zonas mais expostas de litologia de xistos, sem uma ocupação definida à parte de pequenas zonas de cereais. O sector montanhoso permite separar as paisagens mais humanizadas, onde se localizam as aldeias rodeadas por parcelas de culturas variadas para autoconsumo, das zonas superiores, rochosas e de matos, onde pastam os animais, base da economia de montanha.

O agrupamento das variáveis sugere conjuntos de factores relacionados definindo os principais aspectos dos tipos de ambiente de Trás-os-Montes e Alto Douro. A análise de

concentração apresenta uma íntima relação entre os grupos de pontos de referência e os de variáveis, particularmente para os ambientes extremos (vales interiores e montanha atlântica), relação que não é tão nítida nas classes intermédias; pode-se inferir uma maior identidade das regiões que se afastam do padrão dominante da província (planalto xistoso e granítico). A análise de informação mútua do conjunto das variáveis consideradas manifesta a importância da climatologia na sectorização ambiental da província, onde se destacam as variáveis das regiões atlânticas. A importância de cada uma das variáveis na classificação do território sugere o mesmo padrão: as variáveis indicadoras de clima atlântico como as mais importantes, entre as quais se destacam as precipitações mais altas e a nebulosidade. Entre as variáveis fisiográficas destacam-se as que definem a região do vale do Douro (solos antrópicos e socacos estreitos) e entre as de uso do solo, os pastos de montanha e as pastagens húmidas.

Os resultados confrontaram-se com classificações ambientais anteriores, de âmbito nacional e regional. A primeira, de âmbito regional, define Grandes Regiões e foi elaborada para base dos trabalhos de prospecção da Carta de Solos. Os seus sectores relacionam-se bem com os obtidos pela classificação do "TWINSPAN", apesar do seu maior número de classes. Esse maior número de classes é resultado da importância dada às variáveis da litologia, aspecto que se reflecte na análise do valor informativo das variáveis sobre cada um dos sectores considerados; a tal particularidade não será alheia os principais objectivos da sua elaboração.

A segunda cartografia é a Carta Ecológica realizada nos anos 40, com objectivos de zonar o território nacional em sectores homogéneos fito-edafo-climaticamente. A sua sectorização determina um grande número de classes, das quais a maior parte ocupa uma posição intermédia entre os tipos de paisagem definidos pelo "TWINSPAN". Apesar disso, ficam reflectidas as maiores tendências de classificação, gradientes atlânticos e mediterrâneos, estes últimos discriminados pela proximidade ao Vale do Douro. A identificação confirma-se pela importância das variáveis na sectorização realizada; as variáveis climáticas e de uso do solo são praticamente as mesmas da classificação do "TWINSPAN"; as diferenças surgem nas variáveis fisiográficas, nas quais assumem papel preponderante os níveis altitudinais considerados. O resultado afigura-se coerente por se tratar de uma classificação baseada sobretudo na altitude.

A terceira cartografia tem a mesma autoria da anterior, sendo em vários aspectos muito similar. A Carta de Agrotipos é contemporânea da Carta Ecológica e considera aspectos da paisagem agrária na classificação do território. Quando comparada com a do "TWINSPAN", a sua identificação é notável, o que se reflecte na análise canónica de tabela de contingência. A importância das variáveis nesta sectorização segue o padrão da do "TWINSPAN", excepto numa maior importância dos níveis altitudinais, tal como a anterior.

A análise da coincidência geral das cartografias foi realizada pelos índices de Sørensen e Ochiai. Os resultados apresentam uma quase perfeita identidade entre as cartografias Ecológica e de Agrotipos, a que não será alheio o facto de terem o mesmo autor, data e base conceptual. As maiores coincidências, depois da anterior, surgem entre as cartografias das Grandes Regiões com a "TWINSPAN", e de Agrotipos com a de "TWINSPAN"; a coincidência entre as Grandes Regiões e Agrotipos é a de menor valor entre todas, apresentando-se por isso a classificação do "TWINSPAN", como uma classificação intermédia dessas duas.

Conclusão

Como primeiro resultado do estudo, pode-se destacar o definir com clareza a organização funcional da província, do ponto de vista ecológico. Para isso foi fundamental lançar mão de ferramentas modernas de análises de dados, como são os métodos multidimensionais e a teoria da informação. Fica sublinhada a diversidade de gradientes atlânticos e mediterrâneos da província, derivados do cruzamento de factores como um vale muito pronunciado como é o Douro, e acidentes orográficos como o complexo montanhoso galaico-duriense, sobre uma base residual da meseta centro-ibérica. Apresentam-se 8 quadros paisagísticos, sinónimos de funcionamentos muito particulares do meio natural, que se pretende que sejam uma possível base para reflectir sobre o que poderá ser o futuro desenvolvimento de Trás-os-Montes e Alto Douro.

SINTESIS

I - REVISIÓN DE METODOLOGÍAS DE ANÁLISIS DE DATOS

Métodos de análisis multivariante

Las metodologías de análisis multivariante, junto al cálculo matricial, el análisis dimensional y el análisis de la información, son herramientas matemáticas potentes utilizadas hoy en día para la investigación, la comprensión y la modelación de sistemas ecológicos. Su elevado valor se debe a la posibilidad de considerar los sistemas naturales como conjuntos de objetos caracterizados por un elevado número de factores (descriptores), más o menos inter-relacionados entre sí, y con distintos tipos de respuestas a la evolución de esos sistemas.

El desarrollo de las técnicas de análisis multivariante se hace ya desde principios de este siglo, pero el uso de los ordenadores ha sido la causa de que dichas técnicas, en los últimos treinta años, se hayan tornado más accesibles y potentes para el análisis de grandes cantidades de información. Hoy en día se diferencia entre los métodos de clasificación y de ordenación, ambos basados en el análisis de semejanza entre objetos: los primeros sirven para formar grupos de objetos, con una identidad bien definida en función de los respectivos descriptores; los segundos sirven para determinar la organización - disposición relativa - de los objetos, teniendo en cuenta las coincidencias de descriptores en ellos.

Los métodos de clasificación, en función del desarrollo de su algoritmo, se definen como aglomerativos o divisivos; los primeros se aplican a los casos de objetos perfectamente definidos e individualizados, y su procedimiento aglomerativo traduce la evolución común de esos individuos; los divisivos intentan definir fronteras en un universo que se presenta como un *continuum*, determinando subdivisiones con identidad bien marcada. Las clasificaciones se dicen jerárquicas cuando el acercamiento a su resultado se hace de una forma progresiva, de niveles inferiores a superiores (aglomerativas), o vice-versa (divisivas), en la organización de los objetos. Según el número de variables tenidas en cuenta en los procesos clasificativos, las técnicas de clasificación se dicen "monotéticas" o "politéticas".

Los procedimientos aglomerativos se distinguen por lo que determina la fusión de objetos y/o sus grupos; se puede tener en cuenta las "distancias" entre ellos, la homogeneidad de los grupos formados, el aporte de información de la clasificación en su conjunto, etc. De fundamental importancia es la elección del índice que mide la semejanza entre objetos y/o grupos, de acuerdo con la naturaleza de los datos, nominales, ordinales o reales. Los procedimientos divisivos se sirven de la presencia/ausencia de determinadas variables, para repartir el conjunto a clasificar; la elección

de esas variables puede hacerse por criterios de información mútua, valores medios de chi-cuadrado o aún por la preponderancia con que se sitúan en determinado extremo de un eje de ordenación de los objetos a clasificar.

Los métodos de ordenación se utilizan en análisis directos o indirectos de gradientes de datos en estudio. Sus procedimientos se basan en la construcción de una matriz que traduce la semejanza entre los objetos, la cual es transformada por decomposición espectral. Los métodos directos presuponen un conocimiento previo de la estructura de los datos: la previa repartición de los objetos por grupos identificados con algún factor, o la distinción de las variables entre independientes y dependientes. Entre las técnicas más utilizadas están el análisis discriminante y el de las correlaciones canónicas; el primero indica aquellas variables que mejor reconocen las diferencias entre los grupos de objetos previamente definidos; el segundo cuantifica la responsabilidad de un grupo de variables definido como independiente, sobre el otro considerado como dependiente.

Los métodos indirectos son un acercamiento a la estructura organizativa de los datos. Entre las técnicas de análisis indirectas están los componentes principales, factorial, correspondencias y coordenadas principales; sus diferencias residen en el tipo de datos que consideran, la variación de datos que recojen, la respuesta lineal o no entre las variables, o la ordenación simultánea de objetos y variables. Sus resultados indican ejes que son definidos en función de la participación o peso de las variables utilizadas en el análisis, su importancia en la explicación de la variación de los datos y en la ubicación de los objetos en el espacio multidimensional definido por esos ejes.

Teoría de Información

Los conceptos antagónicos de orden y entropía, ya enunciados en la segunda ley de la termodinámica, son hoy en día la base de desarrollo de modernos conceptos ecológicos. Su cuantificación por Shannon (1948, cit. por Legendre y Legendre, 1984b), como el simétrico de la cantidad de información en un mensaje, ha creado la Teoría de Información y ha transformado la entropía en el parámetro de dispersión por excelencia de variables de tipo nominal, como son las que caracterizan el medio natural. Las metodologías de la teoría de la información se aplican hoy a la modelación de sistemas, estimación de diversidad y análisis estadística de complejos de datos ecológicos.

Margalef (1957, cit. por Daget, 1979) conceptualizó la proporción de especies en una muestra de determinada comunidad como un mensaje sobre la estructura y composición de esa comunidad. De este modo, su valor se utiliza hoy para la cuantificación de propiedades estructurales, diversidad, mutualidad y equivocación en los sistemas ecológicos. La cantidad de información compartida e interactiva entre dos o más variables es también una forma de evaluar la variación conjunta de

distribuciones de frecuencias de tipo nominal, lo que expresa la sociabilidad de especies, interdependencia de factores, etc.(Daget y Godron, 1982).

II - ECOLOGÍA DE TRÁS-OS-MONTES E ALTO DOURO

Área de Estudio

Trás-os-Montes y Alto Douro es una provincia del Nordeste de Portugal, prolongación de la meseta central ibérica cortada al Este y Sur por el valle del Duero y al Oeste por las montañas del macizo galaico-duriense. Ribeiro e Lautensach (1987) la describieron como el "Norte Transmontano" un paisaje cerealista entrecortado por afloramientos rocosos y bosquetes de *Quercus pirenaica*. En el siglo XVII, comenzó a utilizarse el tramo del valle del Duero al Sur de la provincia para el cultivo de la vid en pequeñas terrazas, erigiendo uno de los más espectaculares paisajes agrarios del mundo.

Objetivos

La elección de la región de Trás-os Montes e Alto Douro para la realización de este estudio se debe, ante todo, a la disponibilidad de información sobre las variables de su ecología; la cartografía de los suelos de la provincia se terminó justo antes del comienzo del trabajo, recogiendo gran cantidad de elementos biofísicos y bioclimáticos de la provincia. Se pretendió investigar los tipos de ambiente que ocurren en la provincia, así como la capacidad de resolución de un tipo de muestreo regular en líneas paralelas, al ser contrastado con diversas cartografías temáticas existentes. El territorio se entiende como un *continuum* en el cual hay que definir fronteras que permitan su comprensión y caracterización. Con una perspectiva ecológica - ecología del paisaje - , y a la luz de la teoría general de sistemas, podemos considerarlo estructurado jerárquicamente, como un *holon* de sistemas y subsistemas; la sectorización del espacio natural implica la consideración de múltiples factores operativos e interrelacionados, convergentes en la definición de los diversos gradientes que en él ocurren.

Muestreo

Se investigaron 300 puntos de referencia situados sobre 8 ejes de la Red Geodésica Internacional. Los puntos de muestreo se localizaron en cada minuto de la latitud de la provincia, sobre las longitudes 8°00', 7°45', 7°30', 7°15', 7°00', 6°45', 6°30' y 6°15'; los ejes, escogidos por facilidad de la repetitividad del muestreo (referencia geográfica), coincide no obstante a grandes rasgos con una importante fuente de variación bioclimática: el Norte de influencia sobre todo atlántica, y el sur de características climáticas mediterráneas.

La clasificación ecológica que se realiza a partir del muestreo mencionado es entendida como definición, con criterio múltiple, de "tipos de paisaje", a los que se atribuye una potencialidad productiva o de uso característica. Las variables empleadas para esta clasificación se agrupan en climáticas, fisiográficas y de uso de suelo.

Las variables climáticas se obtuvieron de la caracterización realizada en el mapa de suelos, cubriendo aspectos macro y mesoclimáticos. Están así las temperaturas media, mínima y máxima anuales, su amplitud, el número de días de helada, humedad relativa, precipitación, insolación, nubosidad, evaporación y evapotranspiración.

Las variables fisiográficas han sido muestreadas a partir del mapa topográfico, mapa de suelos y mapa geomorfológico, recogiendo en la medida de lo posible todos los aspectos que caracterizan la base física del espacio transmontano. Están en este caso la altitud, el tipo de relieve, la pendiente, la litología, los rasgos geomorfológicos de ámbito local, el tipo afloramiento rocoso, el suelo y la presencia de paredes de piedra de soporte de tierras.

Con las variables de uso del suelo se intenta considerar los diversos modelos de utilización del espacio que han sido elaborados y perfeccionados a lo largo de los tiempos. Se utilizaron las clases de "uso actual del suelo" que aparecen en el mapa de suelos. En él se refieren los cultivos de secano intensivos y extensivos, los cultivos de regadío, los cultivos de olivo, de vid, de almendro, de castaño y otros frutales, los pastos de montaña y pastizales húmedos y de secano, los bosques de robles y rebollo, de alconorque y encina, pinares y matorrales; se ha considerado tanto su ocurrencia como uso dominante (superior al 50% de la área) o usos compartidos (inferiores al 50%) en la unidad cartográfica correspondiente a cada punto de muestreo, pretendiendo con esto incorporar al estudio aspectos complementarios importantes en el entendimiento de la dinámica del paisaje.

Todas las variables consideradas se han transformado en variables de tipo presencia/ausencia, utilizando las clases definidas por los autores de su cartografía de base (Agroconsultores e COBA, 1990; Ferreira, 1981; Instituto Geográfico e Cadastral, *vários*); esta ha sido la forma de considerar a la vez atributos del tipo cuantitativo real (temperatura, altitud, etc.), atributos ordinales (uso de suelo dominante o dominado) y atributos del tipo cualitativo nominal (litología, suelo, etc). Al final se consideraron 54 variables climáticas, 39 fisiográficas y 28 de uso de suelo.

Análisis de la información

Con la metodología de análisis se pretende, además de clasificar los puntos de muestreo, valorar la capacidad descriptiva de este tipo de muestreo, en relación con las sectorizaciones previas. También es posible valorar la información que aporta cada una de las variables a los

sectores o grupos definidos. Se realizaron análisis de clasificación jerárquica divisiva, de concentración por correlaciones canónicas y análisis por la teoría de información.

La clasificación se hizo por el método "TWINSPAN" (Hill, 1979), una técnica jerárquica divisiva y politética, desarrollada para clasificación de inventarios de vegetación, hoy en día con múltiples aplicaciones en ecología y otras ciencias. Su empleo se justifica por la propia naturaleza del objeto de estudio: el territorio. La clasificación por "TWINSPAN" se basa en la división del primer eje de ordenación canónica de los objetos en estudio; dependiendo de la frecuencia con que estos aparecen en las dos mitades del plano factorial, se identifican atributos definidores (exclusivos de uno de los extremos) y preferenciales (tres veces más frecuentes en uno de los lados); tal definición permite la reordenación de los objetos, en función del estado definidor y/o preferencial de los atributos que caracteriza cada uno de ellos y es esa nueva ordenación la que definitivamente define la partición de los objetos. Cada sub-conjunto formado por la división del conjunto inicial es, a su vez, dividido en otros dos de forma análoga a la descrita antes, un proceso que se repite hasta el nivel definido por el investigador. Los grupos de objetos son obtenidos así en función de sus atributos, de forma que los resultados se presentan en una tabla de doble entrada de grupos de objetos por grupos de atributos.

Varios autores comentan las ventajas e inconvenientes de la utilización de "TWINSPAN" (Gauch, 1982; Jongman e col, 1987; Vos y Stortelder (1992)); entre las ventajas es referida la posibilidad de utilizar datos originales, la clasificación integrada de objetos y sus atributos, la ordenación de los grupos (Gauch, 1982). Entre las segundas se cita la linealidad de sus respuestas (que no suele ocurrir en la naturaleza), la sobrevaloración de objetos con poca información y la subestimación de los casos intermedios.

Resultados

Utilizando tres niveles de partición de las localidades del estudio, se definieron 8 grupos representando otros 8 tipos de paisaje de Trás-os-Montes y Alto Douro; los atributos definidores y/o preferenciales correspondientes nos ayudan a entender su identidad. En el primer nivel divisivo se separan los puntos caracterizados por la influencia atlántica de los de influencia mediterránea, discriminados éstos por factores exclusivamente climáticos: amplitud térmica, nubosidad, etc. Entre las variables preferenciales, además de otras climáticas que confirman las anteriores, se presentan "la altitud de 700 metros" y los "cultivos de cereales de secano extensivo" y "olivares" para los sectores mediterráneos, y "pendientes", "zonas rocosas", "cambisuelos poco abundantes" y " pinares" para los sectores atlánticos. El análisis de las variables permite inferir la separación entre la meseta cerealista transmontana, de las motañas y pendientes atlánticas del eje galaico-duriense.

El segundo nivel separa, en el grupo mediterráneo, los valles del Duero y sus afluentes principales, de las altiplanicies más características de la provincia; sus factores definidores y preferenciales son temperatura, evaporación, evapotranspiración y humedad, días de helada, altitud y pendiente, así como los cultivos de la vid y olivar. A un paisaje de temperaturas suaves, pero seco, de pendientes abruptas en las que se cultiva la vid y el olivo, se contraponen el altiplano de climatología más extrema, mayores precipitaciones y menores evapotranspiraciones, de relieve ondulado y cambisuelos. En el grupo atlántico se separan las zonas de montaña de las de valles más o menos estrechos. Son factores de temperatura y precipitación, evapotranspiración y humedad, altitud y matorrales lo que describe la identidad de los nuevos sectores. Las zonas de valle ("ribeira") tienen clima más suave, la montaña es más compleja, abrupta y salvaje (con abundancia de matorrales).

El tercer nivel matiza las características de los grupos antes definidos, individualizando en cada uno de ellos los sectores, mediante un proceso en el que las variables de fisiografía y de uso de suelo van ganando importancia informativa. En el grupo de los valles mediterráneos se separan los valles de los afluentes del valle del Duero; éste tiene clima más suave y la vid es claramente dominante, ubicada en terrazas erigidas para permitir la ocupación de sus abruptas pendientes; los valles de los afluentes tienen relieve ondulado, son más áridos que el anterior y se cultiva preferencialmente cereales de secano extensivos. Del grupo del altiplano se separan sectores que se diferencian por su litología y topografía, un grupo granítico y rocoso, utilizado para vid y pinar, y otro que agrupa la de esquistos junto con las rocas básicas, en los que entremezclan zonas secas de cereales y zonas más húmedas con pastizales de secano. En la parte atlántica, en el grupo de las riberas se separan sus zonas más estrechas y húmedas organizadas en amplias terrazas con cultivos de regadío, de las zonas más expuestas de litología de esquistos, sin una ocupación definida aparte de pequeñas zonas de cereales. El sector montañoso permite separar los paisajes más humanizados, con la ubicación de los pueblos rodeados por parcelas de cultivo heterogéneo de autoconsumo, de las zonas superiores, rocosas y de matorrales, donde pastan los animales que están en la base de la economía de montaña.

El agrupamiento de las variables sugiere conjuntos de factores inter-relacionados definiendo los principales rasgos de los tipos de ambiente de Trás-os-Montes y Alto Douro. El análisis de concentración presenta una íntima relación entre los grupos de puntos de referencia y los de variables para los ambientes extremos (valle del Duero y montaña atlántica), relación esa que no es tan nítida en clases intermedias; se puede inferir una mayor identidad de las regiones que se alejan del patrón dominante de la provincia (altiplano de esquistos y de granito). El análisis de la información mútua al conjunto de factores considerados pone de manifiesto la importancia de la climatología en la sectorización ambiental de la provincia, destacándose las variables de las

regiones atlánticas. La importancia de cada una de las variables en la clasificación del territorio sigue el mismo patrón: primero las variables indicadoras de clima atlántico, donde destacan las altas precipitaciones y nubosidad. Entre las variables fisiográficas destacan a que definen la región del valle del Duero (suelos antrópicos en estrechas terrazas) y a de uso de suelo como pastos de montaña y pastizales húmedos.

Los resultados fueron confrontados con anteriores clasificaciones ambientales, de ámbito nacional o regional. La primera, de ámbito regional, define Grandes Regiones Naturales y se hizo para base de trabajos de prospección de la mapa de suelos. Sus sectores se relacionan bien con los obtenidos por "TWINSPAN", aparte de un mayor número de clases. Ese mayor número de clases resulta de la importancia dada a las variables de litología, aspecto que queda bien reflejado en el análisis al valor informativo de las variables sobre cada uno de los sectores considerados; tal particularidad no será ajena a los objetivos principales de su elaboración.

La segunda cartografía es la Carta Ecológica realizada en los años 40, con objetivos de dividir el territorio nacional en sectores homogéneos de significado fito-edafo-climático. Su sectorización determina un gran número de clases, de las cuales la mayor parte ocupa una posición intermedia entre los tipos de paisaje definidos por "TWINSPAN". Quedan sin embargo reflejadas la identificación de las mayores tendencias de clasificación, gradientes atlánticos y mediterráneos, estos últimos discriminados por la proximidad al valle del Duero. La identificación se confirma por el análisis de la importancia de las variables en la sectorización hecha; las variables climáticas y de uso del suelo son prácticamente las mismas de la clasificación de "TWINSPAN"; las diferencias surgen en las variables fisiográficas, en las que asumen papel preponderante los niveles altitudinales. El resultado es coherente por tratarse de una clasificación principalmente basada en la altitud.

La tercera cartografía tiene la misma autoría de la anterior, siendo en varios aspectos muy similares. La Carta de Agrotipos es contemporánea de la Carta Ecológica y considera sólo aspectos de paisaje agrario a la clasificación del territorio. Cuando se compara con la de "TWINSPAN", su identificación es notable y queda bien reflejada en el análisis canónico de la tabla de contingencia. La importancia de las variables en esta sectorización sigue el patrón de la de "TWINSPAN", aparte de una mayor importancia de los niveles altitudinales, tal como la anterior.

El análisis de la coincidencia general de las cartografías se hizo recurriendo a los índices de disimilitud de Sørensen y Ochiai. Los resultados presentan una casi perfecta identidad entre las cartografías Ecológica y de Agrotipos, a la que no será ajeno el hecho de tener el mismo autor, fecha y base conceptual. Las mayores coincidencias, después de las anteriores, son las cartografías de Regiones Naturales con la de "TWINSPAN", y la de Agrotipos con la de "TWINSPAN"; la

coincidencia entre las Regiones Naturales y Agrotipos es la menor de entre todas, se presentando la clasificación de "TWINSPAN" como una clasificación intermedia de esas dos.

Conclusión

Como primer resultado del estudio, se puede destacar el definir con claridad la organización funcional de la provincia desde el punto de vista ecológico. Para eso ha sido fundamental echar mano de herramientas modernas de análisis de datos como son los métodos multidimensionales y la teoría de información. Queda reflejada la diversidad de gradientes atlánticos y mediterráneos de la provincia, derivados del cruce de factores como un valle muy pronunciado como es el Duero y accidentes orográficos como el complejo montañoso Galaico-duriense, sobre una base residual de la meseta centro-ibérica. Se presentan 8 cuadros paisajísticos sinónimos de funcionamientos muy particulares del medio natural, que se pretende que sean una posible base para reflexionar sobre lo que podrá ser el futuro desarrollo de Trás-os-Montes y Alto Douro.

SYNTHÈSE

I - RÉVISION DE MÉTHODOLOGIES D'ANALYSE DE DONNÉS

Méthodes d'analyse multidimensionnelle

Les méthodologies d'analyse multidimensionnelle, assemblés au calcul matriciel, à l'analyse dimensionnelle et au calcul informationnel, sont des puissants outils mathématiques utilisés aujourd'hui dans l'investigation, la compréhension et la modélisation des systèmes écologiques. Leur grande contribution se doit à la possibilité de considérer les systèmes naturels en tant qu'ensembles d'objets (échantillons, localités) caractérisés par une infinité de descripteurs (variables) plus ou moins reliés entre eux; se sont des systèmes qui présentent des types distincts de réponses à l'évolution des écosystèmes.

Le développement des techniques multidimensionnelles se déroule depuis le commencement du siècle, mais ont été les ordinateurs qui, dans trente dernières années, les ont rendues accessibles et puissantes dans l'analyse de grandes quantités d'information. De nos jours, on fait une différence entre les méthodes de groupement et d'ordonnance, toutes les deux basées sur la ressemblance d'objets: les premières servent à former des groupes d'objets avec une identité qui est fonction des variables respectives; les secondes servent à déterminer l'organisation des objets, ayant en compte les coïncidences des variables en eux.

Les méthodes de groupement se définissent en tant que par agglomération et par division, en fonction du développement de leur algorithme: Les premières s'appliquent aux cas d'objets parfaitement définis et individualisés, et leur procédure par agglomération traduit l'évolution commune de ces individus; les méthodes par division essaient de définir des frontières dans un univers qui se présente comme un *continuum*, en essayant de déterminer des subdivisions avec une identité bien établie. Les groupements se disent hiérarchiques lorsque l'approche à leurs résultats se fait de forme progressive (par agglomération) ou le contraire (par division), dans l'organisation des objets. Selon le nombre de variables considérées dans les procédures classificatrices, les techniques de groupement se disent "monothétiques" ou "polythétiques".

Les procédés agglomérants se distinguent par ce qui détermine la fusion des objets et/ou leurs groupes: on peut prendre en considération les distances entre eux, l'homogénéité des groupes formés ou l'augmentation de l'information dans son ensemble. D'une importance fondamentale est le choix de l'index qui mesure la ressemblance entre objets et/ou groupes, selon la nature des données, nominales, ordinales ou réelles. Les procédés par division se servent de la présence/absence de

certaines variables pour partager l'ensemble à classifier; la sélection de ces variables peut être faite par des critères d'information mutuelle, par des valeurs moyennes de khi carré ou encore par la prépondérance avec laquelle elles ont lieu dans une extrémité donnée d'un axe d'ordination des objets à classifier.

Les méthodes d'ordination s'utilisent dans des analyses directes ou indirectes de gradients des données sous étude. Leurs procédés se fondent sur une matrice qui traduit la ressemblance entre les objets par décomposition spectrale. Les méthodes directes présupposent une connaissance préalable de la structure des données: la répartition préalable des objets parmi des groupes identifiés avec un facteur quelconque, ou la distinction des variables entre dépendantes et indépendantes. Entre les techniques les plus utilisées se trouve l'analyse discriminante et celles des corrélations canoniques: la première indique celles des variables qui reconnaissent le mieux les différences entre les groupes d'objets préalablement définis; la deuxième quantifie la responsabilité d'un groupe de variables définis comme indépendant, sur un autre considéré comme dépendant.

Les méthodes indirectes constituent une approche à la structure organisationnelle des données. Parmi les techniques d'analyse indirecte se trouvent celle des composants principaux, la factorielle, celle des correspondances et coordonnées principales; leurs différences se trouvent dans le type de données qu'elles considèrent, la variation de données qu'elles analysent et l'ordination simultanée d'objets et variables. Leurs résultats indiquent des axes qui sont "participés" par les variables utilisées dans l'analyse, leur importance dans l'explication de la variation des données et la localisation des données dans l'espace multidimensionnel défini par ces axes.

Théorie de l'information

Les concepts antagoniques d'ordre et d'entropie, qui sont depuis longtemps énoncés dans la deuxième loi de la thermodynamique, sont de nos jours la base de développement de concepts écologiques modernes. La quantification par Shannon (1948, cité par Legendre e Legendre, 1984b) de la quantité d'information d'un message, créa la théorie de l'information et transforma l'entropie dans le paramètre de dispersion par excellence de variables du type nominal, comme celles qui caractérisent l'environnement naturel. Les méthodologies de la théorie de l'information s'appliquent aujourd'hui à la modélisation de systèmes, à l'estimation de la diversité et à l'analyse statistique de complexes de données écologiques.

Margalef (1957, cité par Daget, 1979) conceptualisa la proportion des espèces dans un échantillon d'une communauté déterminée comme un message sur la structure et la composition de cette même communauté. De cette façon, cette valeur est utilisée de nos jours pour la quantification

de propriétés structurelles, diversité, mutualité et méprise des systèmes écologiques. La quantité d'information partagée et interactive entre deux ou plus variables est aussi un moyen d'évaluer la co-variation de distributions du type nominal, ce qui exprime la sociabilité des espèces, l'interdépendance de descripteurs, etc. (Daget e Godron, 1982).

II - ECOLOGIE DE TRÁS-OS-MONTES E ALTO DOURO

Région de l'étude

"Trás-os-Montes e Alto Douro", une province du nord-est du Portugal, c'est une extension du plateau ibérique coupée à l'Est et au Sud par la vallée du "Douro" et au Nord et à l'Ouest par le massif "Galaico-Duriense". Ribeiro et Lautensach (1987) l'ont décrite comme le "Norte Transmontano", un paysage de céréales coupé par des affleurements rocheux et des petits bois de *Quercus pirenaica*. Au 17^{ème} siècle, la partie de la vallée du Douro au Sud de la province commença à être utilisée pour la culture de la vigne sur de petites terrasses, donnant origine à un des paysages agricoles les plus spectaculaires du monde.

Objectifs

La sélection de la région "Trás-os-Montes e Alto Douro" pour la réalisation de cette étude se doit, avant tout, à la disponibilité d'information sur les variables de son écologie: la cartographie des sols de la province a pris fin avant même le commencement de ce travail, recueillant beaucoup d'éléments biophysiques et bioclimatiques de la province. On a cherché à poursuivre des investigations les types sur qui se trouvent sur cette province tout aussi bien que la validité de la résolution d'un type d'échantillonnage régulier sur des lignes parallèles, sur les différentes cartographies thématiques existantes. Le territoire se développe comme un *continuum* sur lequel s'inscrivent des frontières pour permettre sa compréhension et caractérisation; la perspective de l'environnement, à la lumière de la théorie générale des systèmes, le considère hiérarchiquement structuré, comme un *holon* de systèmes et sous-systèmes; la sectorisation de l'espace naturel implique la considération de plusieurs descripteurs qui opèrent et qui ont des relations entre eux, et qui convergent dans la définition des différents gradients que l'on peut trouver.

Echantillonnage

On a porté les recherches sur 300 points de référence situés sur 8 axes du réseau géodésique international. Les points d'échantillonnage sont localisés sur chaque minute de latitude de la province, sur les longitudes 8°00', 7° 45', 7° 30', 7° 15', 7° 00', 6° 45', 6° 30' et 6° 15'; les axes,

choisis par sa facilité répétitive de l'échantillonnage (référence géographique), coïncident, cependant, avec les tendances majeures de la variation bioclimatique: le Nord d'influence surtout atlantique et le Sud de caractéristiques climatiques méditerranéennes. Le groupement écologique réalisé à partir de l'échantillonnage mentionné s'entend comme définition avec critère multiple de "types de paysage", auxquels une potentialité productive ou d'usage caractéristique sont attribués. Les variables utilisées pour ce groupement sont groupés parmi climatiques, physiographiques et d'usage du sol.

Les variables climatiques furent obtenues à partir de la caractérisation réalisée dans la Carte des Sols, couvrant des aspects macro et meso-climatiques. Ce sont comprises dans ce cas les températures moyenne, minimale et maximale, leur amplitude, le nombre de jours de gelée, l'humidité relative, la précipitation, l'insolation, la nébulosité, l'évaporation et l'évapotranspiration.

Les variables physiographiques furent sélectionnés à partir de la Carte Topographique, de la Carte des Sols et de la Carte Géomorphologique, recueillant les aspects qui définissent la base physique de l'espace de "Trás-os-Montes et Alto Douro". On trouve dans ce cas l'altitude, le type de relief, le déclin, la lithologie, les aspects géomorphologiques de type local, le type d'affleurement rocheux, le sol et la présence de murs de soutient des terres.

Avec les variables d'utilisation du sol on a cherché à considérer les divers modèles d'utilisation de l'espace, qui ont été élaborés et perfectionnés au fur des années. On a utilisé les classes de "usage actuel des sols" qui se trouvent sur la Carte des Sols. On y fait référence aux cultures de terrain non-arrosé intensives et extensives, les cultures de terrains arrosables, les cultures du olivier, de la vigne, de l'amande, du châtaignier et d'autres arbres de fruit, les pâturages de montagne et les pâturages humides et de terrain non-arrosé, les bois de chênes, de chêne-liège et/ou chêne-vert, forêts de pin et buissons; on a considéré soit leur occurrence dominant (plus de 50% de la surface) ou dominé (moins de 50% de la surface), dans l'unité cartographique correspondante à chaque point de l'échantillonnage, en prétendant ainsi incorporer à cette étude des aspects complémentaires importants dans la compréhension de la dynamique du paysage.

Toutes les variables considérées se sont transformées en variables de type présence/absence, les classes définies par les auteurs de leur cartographie de bases ayant été considérées (Agroconsultores e COBA, 1990; Ferreira, 1981; Instituto Geográfico e Cadastral, *divers*); ça a été la façon de joindre des attributs de type quantitatif réel (température, altitude, etc.), des attributs ordinaux (usage du sol dominant ou dominé) et des attributs du type qualitatif nominal (lithologie sols, etc.). On a 54 variables climatiques, 39 variables physiographiques et 28 variables d'usage du sol.

Analyse de l'information

Avec la méthodologie de l'analyse on a prétendu, en outre le groupement des points d'échantillonnage, valoriser la capacité descriptive de ce groupement, ainsi bien que l'information transmise par chacune des variables considérées aux groupes constitués. On a effectué une analyse de groupement hiérarchique par division, une analyse de concentration par des corrélations canoniques et une analyse par la théorie de l'information. Le groupement a été fait par la méthode TWINSpan (Hill, 1979), une méthode hiérarchique par division et "polythétique" développé pour des inventaires de végétation, qui présente aujourd'hui de multiples applications dans le champ de l'écologie et d'autres sciences naturelles. Son usage est justifié par la nature, elle-même, de l'objet de l'étude: le territoire. Le groupement par le TWINSpan se base sur la division du premier axe d'ordination canonique dans les objets en étude; des attributs identificateurs (exclusifs d'une des extrémités) et préférentiels (3 fois plus fréquentes d'un des cotés) sont identifiés, dépendant de la fréquence avec laquelle ils se font remarquer dans les objets des deux cotés du centre de cette dichotomie; une telle définition permet la réordination des objets en fonction de l'état définitif et/ou préférentiel des attributs qui caractérisent chacun d'entre eux; c'est cette nouvelle ordination qui définit définitivement la partition des objets. Chaque ensemble formé par la division initiale est, à son tour, divisé en deux autres sous-groupes de façon analogue à celle décrite, dans une procédure qui est répétée jusqu'au niveau prescrit par l'utilisateur. Les groupes d'objets formés permettent, en fonction de leurs attributs, le regroupement de ces mêmes attributs. Les résultats sont présentés sur une table de double entrée de groupes d'objets par groupes d'attributs.

Plusieurs auteurs mentionnent les avantages et inconvénients de TWINSpan (Gauch, 1982; Jongman *et al.*, 1987, Vos et Stortelder, 1992); entre les avantages on peut faire référence à la possibilité d'utilisation de données originales, le groupement intégré d'objets et de leurs attributs, l'ordination des groupes (Gauch, 1982); entre les inconvénients on cite la linéarité des réponses (qui d'habitude n'apparaît pas dans la nature), la sur-valorisation d'objets peu informatifs et la sous-estimation de données intermédiaires.

Résultats

En utilisant trois niveaux de participation des localités de l'étude on a défini 8 groupes, représentant 8 gradients de paysage de "Trás-os-Montes e Alto Douro"; les attributs définitifs et/ou préférentiels correspondants aident à comprendre leur identité. Le premier niveau de division sépare les localités d'influence atlantique de ceux d'influence méditerranéenne, discriminés par des descripteurs définitifs exclusivement climatiques et caractéristiques de ces gradients: amplitude thermique, nébulosité, etc.... Les variables préférentielles, en outre d'autres climatiques qui

confirment les variables définissantes, présentent l'altitude moyenne de 700 mètres et les cultures de céréales de terrain non-arrosé extensif et de l'olivier pour les secteurs méditerranéens, et des déclives, affleurements rocheux, cambisols dominés et forêts de pin pour les secteurs atlantiques. L'analyse des variables permet d'inférer la séparation entre le plateau céréalière de "Trás-os-Montes et Alto Douro", et les montagnes et déclives de l'axe Galaico-Duriense. Le deuxième niveau sépare, dans le groupe méditerranéen, les vallées du Douro et des ses affluents principaux, des zones de plateaux plus caractéristiques de la province; ses descripteurs définissants et préférentiels sont la température, l'évaporation, l'évapotranspiration et l'humidité, les jours de gelée, l'altitude et déclive, aussi bien que la culture de la vigne et de l'olivier. A un paysage de climat doux mais sec est confronté le plateau, de climatologie extrême, plus hautes précipitations et moindre évapotranspiration, de relief ondulé et cambisols. Dans le groupe atlantique on sépare les zones de montagne de celles de vallées plus ou moins étroites. L'identité des nouveaux secteurs est décrite par des descripteurs de température et précipitation, évapotranspiration et humidité, altitude et buissons. Les régions de vallée ("ribeira") présentent un climat plus doux, la montagne étant plus dure et sauvage (avec une abondance de buissons).

Le troisième niveau fait remarquer les caractéristiques des groupes décrits plus haut, en individualisant pour chacun d'eux les secteurs, moyennant une procédure dans lequel les variables de physiographie et de sols gagnent de l'importance informative. Dans le groupe des vallées de climat méditerranéen, le Douro est séparé de ses affluents; le Douro présente un climat plus doux, la vigne y est nettement dominante, localisée dans des terrasses élevées pour permettre l'occupation de ses pentes rapides; dans les vallées des affluents, plus arides et de relief ondulé, sont cultivés de préférence des céréales de terrains non-arrosés extensifs. Dans le groupe de plateau on sépare des secteurs qui se distinguent par leur lithologie et topographie: un groupe granitique et rocheux, utilisé pour la vigne et le pin, et un autre groupe de schistes et de roches basiques, avec régions sèches de céréales et zones plus humides avec des pâturages de terrain non-arrosé. Dans la région Atlantique, sur le groupe des "ribeiras" on sépare ses zones plus étroites et humides, organisées en de larges terrasses avec des cultures de terrains arrosables, des zones plus exposées présentant une lithologie de schistes, sans occupation définie, à part de petites zones de céréales. Le secteur de montagne permet de séparer les paysages plus humanisés, où se trouvent les villages entourés de parcelles de cultures diverses destinées à l'autoconsommation, des zones supérieures rocheuses et de buissons, où broutent les animaux base de l'économie de montagne.

Le groupement des variables suggère des ensembles de descripteurs en relation entre eux, définissant les aspects principaux des types d'environnements de "Trás-os-Montes e Alto Douro". L'analyse de concentration présente une relation intime entre les groupes de points de référence et les variables pour les ambiances extrêmes (vallées intérieures et montagne atlantique), cette relation

n'étant pas si nette dans les classes intermédiaires; on peut inférer une identité plus grande dans les régions qui s'éloignent de l'étalon dominant de la province (plateau de schiste et granite). L'analyse de l'information mutuelle de l'ensemble de variables considérées manifeste l'importance de la climatologie sur la sectorisation des environnements de la province, où se détachent les variables des régions atlantiques. L'importance de chacune des variables dans la classification du territoire suit le même patron: d'abord les variables climatiques atlantiques parmi lesquelles se détachent les précipitations plus fortes et la nébulosité. Entre les variables physiographiques se détachent celles qui définissent la région de la vallée du Douro (sols anthropiques et terrasses étroites). Entre les variables de l'usage du sol se détachent, les pâturages de montagne et de zones humides.

Les résultats ont été comparés avec d'anciennes classifications de l'environnement, de champ d'action régional et national. La première définit des régions naturelles et a été élaborée en tant que base des travaux de prospection de la Carte des Sols. Ses secteurs entrent assez bien en relation avec ceux obtenus par le groupement TWINSpan, bien que présentant un plus grand nombre de classes. Ce plus grand nombre de classes est le résultat de l'importance accordée aux variables de la lithologie, aspect qui se réfléchit sur l'analyse de la valeur informative des variables sur chacun de secteurs considérés; telle particularité ne sera pas étrangère aux objectifs principaux de son élaboration.

La deuxième cartographie est la Carte Ecologique réalisée dans les années 40 pour repartir le territoire national en des secteurs de caractéristiques phyto-edafo-climatiques homogènes. Sa sectorisation détermine un grand nombre de classes, desquelles la plus grande partie occupe une position intermédiaire parmi les types de paysage définis par le TWINSpan. Malgré ceci, les plus importantes tendances de groupement, gradients atlantiques et méditerranéens, sont réfléchies, les derniers discriminés par la proximité de la vallée du Douro. L'identification est confirmée par l'importance des variables sur la sectorisation effectuée; les variables climatiques et de l'usage du sol plus importantes sont pratiquement les mêmes du groupement TWINSpan; les différences apparaissent dans les variables physiographiques, sur lesquelles exercent un rôle prépondérant les niveaux d'altitude considérés. Ceci constitue un résultat espéré puisqu'il s'agit d'une classification basée surtout sur l'altitude.

La troisième cartographie a les mêmes auteurs que cette dernière, étant en plusieurs aspects très similaires. La Carte d'Agrotypes est contemporaine de la Carte Ecologique et essaye d'incorporer des valeurs de paysage agricole à la classification du territoire. Lorsqu'on la compare avec celle du TWINSpan, son identification est remarquable, ce qui se réfléchit sur l'analyse canonique de la table de contingence. L'importance des variables sur cette sectorisation suit l'étalon de celle du TWINSpan, en faisant exception d'une plus grande importance accordée aux niveaux d'altitude, tout comme la précédente.

L'analyse de la coïncidence générale des cartographies a été réalisée à partir des indexes de Sørensen and Ochiai. Les résultats présentent une identité presque parfaite entre les cartographies Ecologique et d'Agrotypes, phénomène auquel n'est pas étrange le fait qu'elles ont le même auteur, date et base conceptuelle. Les coïncidences les plus grandes, après l'antérieure, sont les cartographies des régions naturelles avec celle du TWISNPAN, et celles d'Agrotype et TWINSPAN; la coïncidence entre les régions naturelles et celles de la Carte d'Agrotypes est la moins importante parmi toutes, le groupement TWISNPAN se présentant comme un groupement intermédiaire entre ces deux-là.

Conclusion

Comme premier résultat de cette étude, on peut faire remarquer la définition claire de l'organisation fonctionnelle de la province, du point de vue écologique. Pour y parvenir ce fût fondamental de faire recours à de modernes outils d'analyse de données, tels que les méthodes multidimensionnelles et la théorie de l'information. La diversité de gradients atlantiques et méditerranéens de la province provenant du croisement de facteurs tels que l'existence d'une vallée très prononcée (Douro) et autres accidents orographiques (montagnes "Galaico-Durienses") sur la base résiduelle du plateau centre-ibérique, est soulignée. On présent huit modèles de paysage, synonymes de fonctionnements très particuliers de l'environnement naturel, qu'on prétende puissent devenir une base possible pour la réflexion sur ce qui pourra être le développement futur de "Trás-os-Montes e Alto Douro".

SYNTHESIS

I - BRIEF DESCRIPTION OF DATA ANALYSIS METHODOLOGIES

Methods of multivariate analysis

The methodologies of multivariate analysis, together with matrix calculus, dimensional analysis and information analysis, are powerful mathematics tools used nowadays in the investigation and understanding and modelisation of ecological systems. Their large contribution is due to the possibility of considering natural systems as sets of objects characterized by an infinity of factors (variables) more or less interconnected among themselves, and with different types of answers to the evolution of ecosystems.

The development of the methodologies of multivariate analysis has been taking place since the beginning of the century, but they were made readily accessible and powerful in the analysis of large quantities of information by computers over the last thirty years. In our days a distinction is made between methods of clustering and ordination, both based on similarity of objects: the former ones are used to form sets of objects with an identity which is function of the respective variables; the latter are used to determine the organization of objects bearing in mind the coincidence of the variables in them.

The methods of clustering in function of the development of their algorithm are defined as agglomeratives and divisives: the first ones apply to cases of perfectly defined and individualized objects and their agglomerative procedure expresses the common evolution of these individuals; divisives methods try to define borders in an universe which presents itself as a *continuum*, aiming at determining subdivisions with a well defined identity. Clusterings are said hierarchical when the approach to their result is performed in a progressive way, from lower to higher levels (agglomerative) or vice-versa (divisive) in the organization of objects. According to the number of variables one bears into account in clustering procedures the techniques are said to be monothetical or polythetic.

Agglomerative procedures are distinguished by what determines the fusion of objects and/or their sets: one might take into account the distances between them, the homogeneity of the sets formed, the increase of clustering information as a whole and so on. Of prime importance is the choice of the index which measures the resemblance among objects and/or sets according to the nature of data, nominal, ordinal, or real. Divisive procedures use the presence/absence of certain variables, to divide the set under clustering: selection of these variables may be done by criteria of

mutual information, chi-square medium values or yet by the preponderance with which they occur in a certain end of an ordination axis of the objects to classify.

Ordination methods are used in direct or indirect gradient analysis of the data under study. Their procedures are based in the construction of a matrix conveying the resemblance among the objects, which is transformed through spectral decomposition. Direct methods presuppose a previous knowledge of data structure: previous distribution of the objects among sets identified with some factor, or the distinction of variables between dependent and independent. Among the most widely used techniques we find discriminant analysis and canonical correlation: the first one indicates those variables which better recognize the differences among the sets of objects previously defined; the second one quantifies the responsibility of a set of variables defined as independent, over another considered as dependent.

Indirect methods are an approach to organizational data structure. Among the indirect analysis techniques are those of principal components, factorial, correspondence and principal coordinates; their differences lay in the sort of data they consider, the variance of data they analyze and the simultaneous ordination of objects and variables. Their results indicate axis which are "participated" by the variables used in the analysis, their importance in the explanation of the variance of data and the location of objects in the multidimensional space defined by those axis.

Theory of information

The antagonistic concepts of order and entropy, which for long have been stated in the second law of thermodynamics, are today the basis for modern ecological concepts. Quantification by Shannon (1948, quoted by Legendre e Legendre, 1984b) of the information of a message, has created the theory of information and has changed entropy in the most perfect parameter of dispersion of variables of a nominal type, like the ones which define natural environment. The methodologies of the theory of information are applied nowadays to the modelisation of systems, estimation of diversity and statistical analysis of clusters of ecological data.

Margalef (1957, quoted by Daget, 1979) has conceptually the proportion of species in a sample of a certain community as a message on the structure and composition of that community. This way its value is used today for the quantification of structural properties, diversity, mutuality and ambiguities of ecological systems. The quantity of information shared and interactive among two or more variables is also a means of evaluating the conjoined variation of frequency distributions of nominal type, which expresses sociability of species, factor interdependence, etc. (Daget e Godron, 1982).

II - ECOLOGY OF TRÁS-OS-MONTES E ALTO DOURO

Area of study

"Trás-os-Montes and Alto Douro", an administrative region in the North-East of Portugal, an extension of the Iberian plateau which is cut at the East and South by the Douro Valley and at the West by the Mountains of the "Galaico-Duriense" Massif. Ribeiro and Lautensach (1987) have described it as the "Norte Transmontano", a cereal landscape penetrate by rock outcrop and small woods of *Quercus pyrenaica*. In the 17th century, the part of the "Douro" valley to the South of the district began to be used to plant vineyards in small terraces, giving birth to one of the most spectacular agricultural landscapes of the world.

Objectives

The selection of the "Trás-os-Montes e Alto Douro" area for the undertaking of this study is due, first and foremost, to the availability of information about the variables of its ecology: the cartography of the soils of the district was done even before the beginning of this study, gathering a large quantity of biophysics and bioclimatics elements of the same district. It was aimed to investigate the different types of environment which can be found in the district, as well as the validity of a type of systematic sampling along parallel lines, over the several existing thematic cartographies. The territory develops as a *continuum* in which borders are determined to allow its understanding and characterization; the environmental perspective, under the light of the general theory of systems considers it hierarchically structured, as an *holon* of systems and subsystems; The partition of the natural space implies the consideration of multiple operative and interconnected factors, converging towards the definition of the different occurring gradients.

Sampling

300 reference points situated on the international geodesic net were investigated. Sampling points were located in each minute of the district's latitude, over the longitudes 8°00', 7°45', 7°30', 7°15', 7°00', 6°45', 6°30' and 6°15'; the axis, chosen according to the ease of recurring frequency of the sampling (geographical reference), coincide, however, with the largest tendencies of bioclimatic variation: the north of mostly Atlantic influence and the South of Mediterranean climatic characteristics. The ecological clustering performed on the basis of the above mentioned sampling is understood as definition with multiple criteria of "types of landscape", to which productive potential or of typical land use is attributed. These variables used are grouped in climatic, physiographic, and of land use.

Climatic variables were obtained from description made in the Soil Map, covering macro and mesoclimatic aspects. In this case we find the average, maximum and minimum temperatures, their

range, the number of days of frost, relative humidity, precipitation, evaporation and evapotranspiration.

Physiographic variables were selected from the Topographic, Soils and Geomorphologic Maps, gathering, as far as possible, all the aspects which define the physical basis of the North-East of "Trás-os-Montes e Alto Douro". In this case are altitude, type of relief, slope degree, lithology, local geomorphologic aspects, type of rock outcrop, soil and the presence of terraces.

With the land use variable it has been tried to consider the different models of space use, which have been built and improved over the times. The classes of "present use of soil" appearing in the Soils Map were used. There, extensive and intensive dry land cultivations, olive, almond, vineyards, chestnut-tree and other fruit trees, mountain pastures, wet and dry pastures, oak woods, cork-oak and/or holm-oak groves, pine woods and bushes are referred; both their occurrence and dominant (over 50% of the area) or dominated (under 50% of the area) use were considered in the cartographic unit corresponding to each point of the sampling, thus aiming at blending in the study important complementary aspects in the understanding of landscape dynamics.

All the variables considered were transformed into variables of the absence/presence kind, the classes used by the authors of their basic cartography, having been used (Agroconsultores e COBA, 1990; Ferreira, 1981; Instituto Geográfico e Cadastral, *various*); this was the way of joining together attributes of a real quantitative type (temperature, altitude, etc.), ordinal attributes (dominant or dominate land use) and attributes of the qualitative nominal type (lithology, soils, etc.). 54 climatic variables, 39 physiographic variables and 28 variables of land use were considered.

Analysis of information

With the methodology of analysis we have aimed at, apart from classifying the sampling points, to valorize the descriptive capacity of that classification, and the information conveyed by each of the considered variables to the constituted sets. Divisive hierarchic clustering analysis, canonical correlation analysis of concentration and analysis using the theory of information were performed. Clustering was made using the TWINSpan method (Hill, 1979), an hierarchic, divisive and polythetical technique, developed for the clustering of vegetation inventory, nowadays presenting multiple uses in ecology and other natural sciences. Its use is justified by the very nature of the object of the study: the territory. TWINSpan clustering is based in the division of the first canonical ordination axis in the objects under study; depending on the frequency with which they occur in the objects at both sides of the centre of the dichotomy, defining attributes (exclusive of one of the extremes) and preferential ones (3 times more frequent on one of the sides) are identified;

such a definition allows the reordering of objects, in function of the defining and/or preferential state of the attributes which define each one of them; it is this new ordering the one which definitely defines the partition of the objects. Each set formed by the initial division is, in its turn, divide into another two by an analogous operation, a procedure which is repeated until the level defined by the user. The groups of objects formed allow, according to their attributes, their grouping. The results are presented in a two-way ordered table of groups of objects by groups of attributes.

Several authors mention the advantages and disadvantages of TWINSpan (Gauch, 1982; Jongman *et al.*, 1987, Vos & Stortelder, 1992); among the advantages, the possibility of the use of original data is mentioned as well as the integrated clustering of objects and their attributes and the ordering of groups (Gauch, 1982); among the disadvantages, the linearity of answers (which usually does not occur in nature) is quoted together with the over-valorization of not very informative objects and the over-estimation of intermediate data.

Results

Using three levels of participation of the localities under study, 8 groups were defined representing 8 landscape gradients of "Trás-os-Montes e Alto Douro"; the defining and/or preferential corresponding attributes help to understand their identity. The first divisive level separates the points of reference of Atlantic influence from those of Mediterranean influence, discriminated by exclusively climatic defining factors and characteristic of those gradients: temperature range, cloudiness, etc. The preferential variables, beyond other climatic which confirm the defining ones, present the average altitude of 700 meters and the extensive dry land cultures of cereal and olive-trees for the Mediterranean sectors, and slopes, rock outcrops, dominate cambisols and pine-woods for the Atlantic sectors. The analysis of variables allows to infer the separation between the cereal plateau from the North-East of "Trás-os-Montes e Alto Douro" and the mountains and slopes of the "galaico-duriense" axis. The second level separates, in the Mediterranean group, the valleys of the Douro and its main affluents, from the plateau areas, more characteristic of the district; its defining and preferential factors are temperature, evaporation, evapotranspiration and humidity, days of frost, altitude and slope, as well as the presence of olive-trees and vineyards. To a landscape with a mild but dry climate, with step slopes where vine and olive are grown, is set against the plateau of a more extreme climatology, bigger local precipitation and less evapotranspiration, with a wavy relief and cambisols. In the Atlantic group the mountain areas are separated from the more or less narrow valley areas. The identity of the new sectors is described through factors of temperature and precipitation, evapotranspiration and humidity, altitude and bushes. The valley areas ("ribeira") have a milder climate, the mountain is tougher and wilder (with abundant bushes).

The third level emphasizes the characteristics of the previously defined groups, individualizing in each of them the sectors, according to a process in which the physiographic and soil variables gain an increasing informative importance. In the Mediterranean valleys group, the "Douro" is separated from its affluents; the Douro has a milder climate, vineyards, located in terraces (ledges) raised to allow the occupation of its steep slopes, are clearly dominant; the valleys of the affluents, of a wavy relief and dryer than the previous one, extensive cereal dry land culture, is preferentially used. From the plateau group, sectors which distinguish themselves by their lithology and topography are separated, a granite rocky group used for vineyards and pine-woods, and another one grouping schist and basic rocks, in which dry areas for cereals are mixed with wetter pasture areas. In the Atlantic area, in the group of the "ribeira" its narrower and wetter areas, organized in wide terraces with wet land cultures, are separated from the more exposed areas, presenting a schist lithology, without a defined occupation, apart from small cereal areas. The mountain sector allows the separation of more humanized landscapes, where villages surrounded by diversified cultures for self-consumption are located, from the higher areas, rocky and bushy, where animals graze, basis of the mountain economy.

The grouping of the variables suggests sets of interconnected factors defining the main aspects of the types of environment of "Trás-os-Montes e Alto Douro". The concentration analysis presents an intimate relationship between the groups of reference points and the variables for extreme environments (inner valleys and Atlantic mountain), such a relationship not being so clear in the intermediate classes; a bigger identity which differ from the dominating pattern of the district may be inferred (schist and granite plateau). The analysis of the mutual information of the set of variables considered shows the importance of climatology in the environmental division of the district, where the variables of the Atlantic areas are outstanding. The importance of each one of the variables in the classification of the territory follows the same pattern: first the Atlantic climatic variables, among which nebulosity and higher precipitation are outstanding. Among the physiographic variables those defining the area of the "Douro" valley (anthropic soils and narrow terraces) and those of the land use in mountain pastures and wet landscapes are outstanding.

The results were collated with older environmental classifications, of national or regional type. The first one, of a regional character, defines Natural Regions and was prepared having the prospecting works for the Soils Map as basis. Its sectors interconnect well with those obtained from the TWINSPAN classification, although presenting a bigger number of classes. This bigger number of classes is the result of the importance given to the lithology variables, aspect which reflects on the analysis of the informative value of the variables on each of the considered sectors; such a peculiarity is certainly not alien to the main goals of its elaboration.

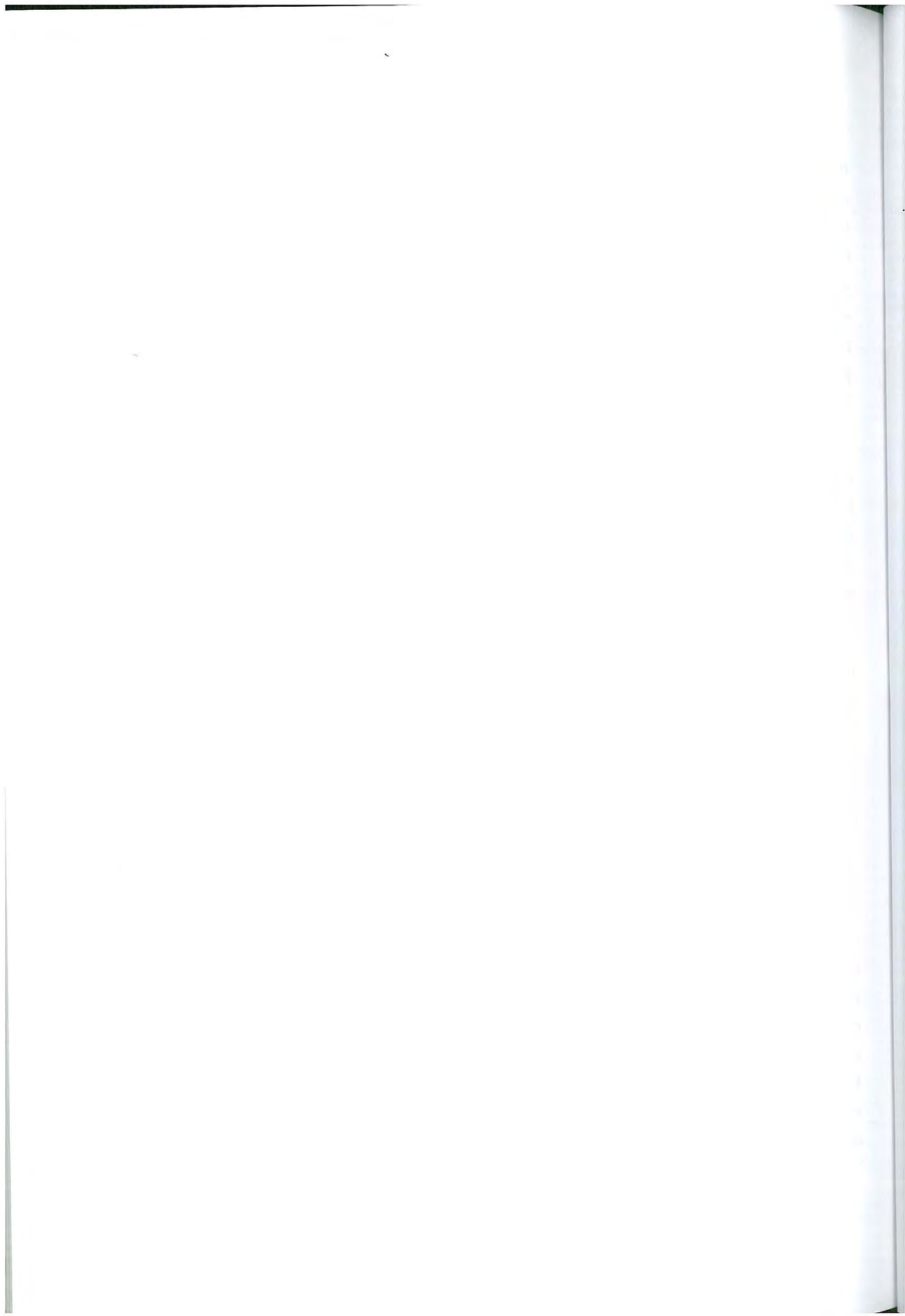
The second cartography is the Ecological Map accomplished during the forties, with the aim of dividing the national territory in phyto-edaphic-climatic homogeneous sectors. Its partition determines a large number of classes, from which a large part occupies an intermediate position among the types of landscape defined by TWINSpan. In spite of this, the major trends of classification, Atlantic and Mediterranean gradients are defined, these last ones discriminated according to their proximity to the Douro Valley. The identification is confirmed by the importance of the variables in the partition which took place; the climatic variables and of land use are practically the same as those from the TWINSpan classification; the differences appear in the physiographic variables, in which the levels of altitude considered assume an outstanding role. This would have been expected, since it is a classification, based mostly on altitude.

The third cartography has the same authorship of the previous one, having in many aspects a large resemblance. The Map of Agrotypes is contemporary of the Ecological Map and tries to incorporate landscape agricultural values to the classification of the territory. When compared to the TWINSpan one, its identification is remarkable, which is reflected in the canonical analysis of the contingency table. The importance of the variables in this partition follows the pattern of the one from TWINSpan, except for a larger importance for the altitude levels, like the preceding one.

The analysis of general coincidence of the cartographies was done by the Sørensen and Ochiai indexes. The results present an almost perfect identity between the Ecological and Agrotypes cartographies, to which the fact that both have the same author, date and conceptual basis shall not be alien. The major coincidences, apart from the former one, are those of the cartographies of Natural Regions with those from TWINSpan, and the Agrotypes and TWINSpan; the coincidence between Natural Regions and Agrotypes is the less important one, the TWINSpan classification presented as an intermediate classification between those two.

Conclusion

As a first result of the study one can highlight the clear definition of the functional organization of the district, on an ecological point of view. To achieve this it was fundamental to use modern tools for data analysis, such as multidimensional methods and the theory of information. The diversity of Atlantic and Mediterranean gradients of the district, resulting from the crossing of factors as presence of a very pronounced valley such as the Douro and orographic accidents as the mountain system on a residual base of the central-Iberian plateau, was underlined. Eight landscape tables are presented, synonymous of very particular operation of the natural environment, which we aim will become a possible basis for reflection on what might be the development of "Trás-os-Montes e Alto Douro".



BIBLIOGRAFIA

- AGROCONSULTORES E COBA **1987** - Carta dos solos e carta de utilização actual do solo do nordeste de Portugal: 1ª versão preliminar. UTAD, Vila Real.
- AGROCONSULTORES E COBA **1990** - Carta dos solos, carta de utilização actual do solo e carta de uso potencial do solo do nordeste de Portugal: 2ª versão preliminar. UTAD, Vila Real.
- ALBUQUERQUE, J. DE PINA MANIQUE E **1945** - Zonas fito-climáticas e regiões naturais do continente português. *Boletim da Sociedade Broteriana*, XIX (2ª série), p. 569-591.
- ALBUQUERQUE, J. DE PINA MANIQUE E **1954** - Carta Ecológica de Portugal. Ministério da Economia, Lisboa.
- ALVAREZ, J., YANGUAS, M. M., RODRIGUEZ, M. A., GÓMEZ SAL, A. **1989** - Condicionantes geofísicos de los usos del territorio en un valle de la montaña cantábrica. *in* Options Méditerranéennes - Série Séminaires N° 3, p. 141-152.
- CICÉRI, M.-F., MARCHAND, P., RIMBERT, S. **1977** - Introduction à l'analyse de l'espace. Masson, Paris.
- DAGET, J. **1979** - Les Modèles Mathématiques en Écologie. Masson, Paris, 2ª Edição.
- DAGET, PH e GODRON, M. **1982** - Analyse de l'écologie des espèces dans les communautés. Masson, Paris.
- DE PABLO, C. L., GÓMEZ SAL, A., PINEDA, F. D. **1987** - Élaboration automatique d'une cartographie écologique et son évaluation avec des paramètres de la théorie de l'information. *in* L'Espace Géographique, N° 2, p. 115-128.
- DE PABLO, C. L., MARTÍN DE AGAR, P., GÓMEZ SAL, A., PINEDA, F. D. **1988** - Descriptive capacity and indicative value of territorial variables in ecological cartography. *in* Landscape Ecology, N° 1(4), p. 203-211.
- DIGBY, P. e KEMPTON, R. **1987** - Multivariate Analysis of Ecological Communities. Chapman and Hall, London.
- FEOLI, E. e ORLÓCI, L. - **1979** - Analysis of concentration and detection of underlying factors in structured tables. *Vegetatio*, 40, p. 49-54.

- FERREIRA, Denise de Brum **1981** - Carte Geomorphologique du Portugal. INIC, Lisboa.
- FRONTIER, S. **1982** - Stratégies d'Échantillonnage en Écologie. Masson, Paris.
- FRY, J. C. (EDS.) **1983** - Biological Data Analysis - A Practical Approach. Oxford University Press, Oxford.
- GAUCH JR., HUGH G. **1982** - Multivariate Analysis in Community Ecology. Cambridge University Press, Cambridge.
- GAUCH, H.G. e WHITTAKER, R.H. **1981** - Hierarchical classification of community data. *Journal of Ecology*, 69, p. 537-557.
- GÓMEZ SAL, A. e MUÑOZ-YANGUAS, M. A. **1990** - Análisis de la coincidencia entre componentes del paisaje mediante parámetros de información. I Congreso de Ciencia del Paisaje, Barcelona. Monografies de L'Equip 3. p. 223-242.
- GÓMEZ SAL, A., MUÑOZ-YANGUAS, M. A., REBOLLO, S. **1990** - Valoración del paisaje de puertos de montaña en función de sus facies bióticas e geológicas. I Congreso de Ciencia del Paisaje, Barcelona. Monografies de L'Equip 3. p. 243-255.
- GÓMEZ SAL, A., MUÑOZ-YANGUAS, M. A., SAENZ DE BURUAGA, M., COSTA, L., REBOLLO, S. **1992** - Relaciones entre valor natural y complejidad estructural en un territorio de la Cordillera Cantábrica. in DÍAZ PINEDA, F., CASADO, M. A., DE MIGUEL, J. M., MONTALVO, J. (eds.). Diversidade Biológica. Fundación Ramón Areces. Madrid. p. 71-77.
- GÓMEZ SAL, A., ALVAREZ, J., MUÑOZ-YANGUAS, M. A., REBOLLO, S. **1993** - Patterns of Change in the Agrarian Landscape in an Area of the Cantabrian Mountains (Spain) - Assessment by Transition Probabilities. in BUNCE, R. G. H., RYSZKOWSKI, L., PAOLETTI, M. G. (eds.). Landscape Ecology and Agroecosystems. Lewis Publishers. Boca Raton. p. 141-152.
- GRUPO DE TRABALHO DO ATLAS DO AMBIENTE **1978** - Carta da distribuição de Carvalhos e Castanheiro II.3 in Atlas do Ambiente Escala 1:1 000 000. Comissão Nacional do Ambiente/Secretaria de Estado do Ambiente e Recursos Naturais, Lisboa.
- GRUPO DE TRABALHO DO ATLAS DO AMBIENTE **1982** - Carta Hipsométrica I.15 in Atlas do Ambiente Escala 1:1 000 000. Comissão Nacional do Ambiente/Secretaria de Estado do Ambiente e Recursos Naturais, Lisboa.

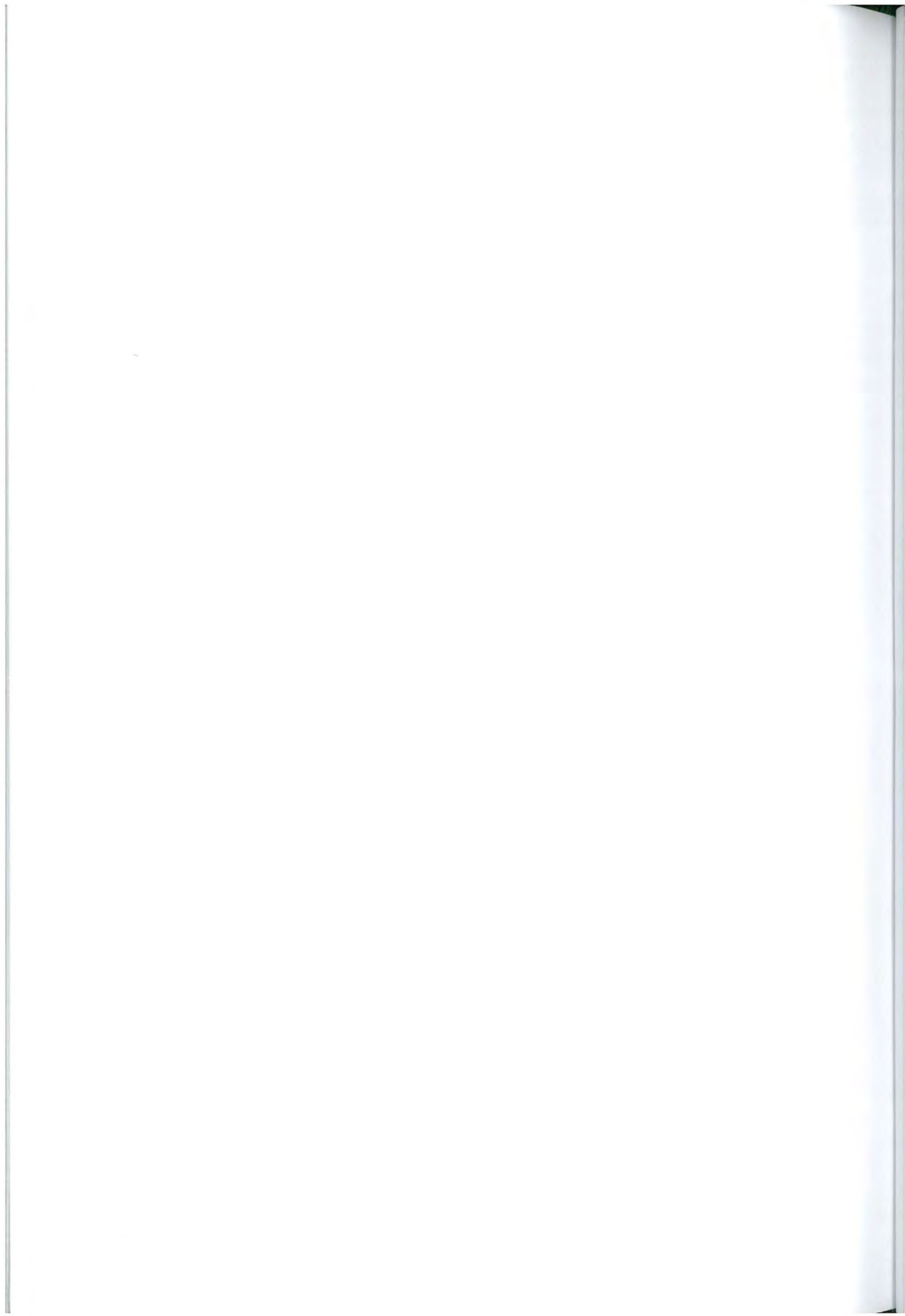
- GRUPO DE TRABALHO DO ATLAS DO AMBIENTE **1984** - Carta Ecológica Fito-edafo-climática III.4 *in* Atlas do Ambiente Escala 1:1 000 000. Comissão Nacional do Ambiente/Secretaria de Estado do Ambiente e Recursos Naturais, Lisboa.
- GRUPO DE TRABALHO DO ATLAS DO AMBIENTE **1985** - Carta da distribuição de culturas arvenses de sequeiro II.15 *in* Atlas do Ambiente Escala 1:1 000 000. Comissão Nacional do Ambiente/Secretaria de Estado do Ambiente e Recursos Naturais, Lisboa.
- GRUPO DE TRABALHO DO ATLAS DO AMBIENTE **1985** - Regiões Naturais Caracterização Eco-fisionómica III.5 *in* Atlas do Ambiente Escala 1:1 000 000. Comissão Nacional do Ambiente/Secretaria de Estado do Ambiente e Recursos Naturais, Lisboa.
- HILL, M. O. **1979** - TWINSPAN - A FORTRAN Program for Arranging Multivariate Data in an Ordered Two-way Table by Classification of Individuals and Attributes. Cornell University, Ithaca, New York.
- HILL, M.O., BUNCE, R.G.H., SHAM, M.W. **1975** - Indicator species analysis, a divisive polythetic method of classification, and its application to a survey of native pinewoods in Scotland. *Journal of Ecology*, 63, p. 597-613.
- INSTITUTO GEOGRÁFICO E CADASTRAL **1962** - Carta de Portugal 1:100 000 Folha 2 Portos-Tourém. IGCP, Lisboa.
- INSTITUTO GEOGRÁFICO E CADASTRAL **1964** - Carta de Portugal 1:100 000 Folha 12 Bemposta. IGCP, Lisboa.
- INSTITUTO GEOGRÁFICO E CADASTRAL **1965** - Carta de Portugal 1:100 000 Folha 3 Vinhais. IGCP, Lisboa.
- INSTITUTO GEOGRÁFICO E CADASTRAL **1966** - Carta de Portugal 1:100 000 Folha 11 Torre de Moncorvo. IGCP, Lisboa.
- INSTITUTO GEOGRÁFICO E CADASTRAL **1970a** - Carta de Portugal 1:100 000 Folha 10 Vila Real. IGCP, Lisboa.
- INSTITUTO GEOGRÁFICO E CADASTRAL **1970b** - Carta de Portugal 1:100 000 Folha 6 Chaves. IGCP, Lisboa.
- INSTITUTO GEOGRÁFICO E CADASTRAL **1971** - Carta de Portugal 1:100 000 Folha 14 Lamego. IGCP, Lisboa.

- INSTITUTO GEOGRÁFICO E CADASTRAL **1973** - Carta de Portugal 1:100 000 Folha 15 Pinhel. IGCP, Lisboa.
- INSTITUTO GEOGRÁFICO E CADASTRAL **1974** - Carta de Portugal 1:100 000 Folha 8 Miranda do Douro. IGCP, Lisboa.
- INSTITUTO GEOGRÁFICO E CADASTRAL **1990** - Carta de Portugal 1:100 000 Folha 4 Deilão. IGCP, Lisboa.
- INSTITUTO GEOGRÁFICO E CADASTRAL **1992** - Carta de Portugal 1:100 000 Folha 7 Bragança. IGCP, Lisboa.
- JONGMAN, R. H., TER BRAAK, C. J. F., VAN TONGEREN, O. F. R. **1987** - Data analysis in community and landscape ecology . Pudoc Scientific Publishers, Wageningen.
- LEECH, J. W., CORRELL, R. L., AUNG KYAW MYINT **1991** - Use of principal-coordinate analysis to assist in aggregating species for volume-table construction. *Forest Ecology and Management*, 40, 279-288.
- LEFEBVRE, J. **1983** - Introduction aux Analyses Statistiques Multidimensionnelles. Masson, Paris, 3ª Edição
- LEGENDRE, L. e LEGENDRE , P. **1984a** - Écologie numérique 1 .Le traitement multiple des données écologiques. Masson, Paris, 2ª Edição.
- LEGENDRE , L. e LEGENDRE , P. **1984b** - Écologie numérique 2 .La structure des données écologiques. Masson, Paris, 2ª Edição.
- LEPART, J. e DEBUSSCHE, M. **1980** - Information efficiency and regional constellation of environmental variables. *Vegetatio*, 42, 85-91.
- LUDWIG, J. e REYNOLDS, J. **1988** - Statistical Ecology - A primer methods and computing. John Wiley e Sons, New York.
- MOPT **1992** - Guía para la elaboración de estudios del medio físico. *Serie Monografías*. Ministerio de Obras Públicas y Transportes, Madrid.
- NAVEH, Z. e LIEBERMAN, A. **1983** - Landscape Ecology - Theory and Application. Springer-Verlag, New York.

- ORLÓCI, L. **1969** - Information analysis of structure in biological collections. *Nature*, 223, p. 483-484.
- ORLÓCI, L. **1976** - Ranking characters by an information criterion. *Journal of Ecology*, 64, p. 417-419.
- ORLÓCI, L. **1991a** - Entropy and Information. SPB Academic Publishing bv, The Hague.
- ORLÓCI, L. **1991b** - CONAPACK: Program for Canonical Analysis of Classification Tables. Ecological Computations Series (ECS): Vol. 4. SPB Academic Publishing bv, The Hague.
- PODANI, J. **1979** - Association-analysis based on the use of mutual information. *Acta Botanica Hungarica*, 25, p. 125-130.
- PODANI, J. **1989** - Comparison of ordinations and classifications of vegetation data. *Vegetatio*, 83, p. 111-128.
- PODANI, J. **1990** - SYN-TAX IV Computer Programs for Analysis in Ecology and Systematics on IBM-PC and Macintosh Computers. UNIDO - ICS, New York.
- PODANI, J. **1993** - SYN-TAX-5.0: Computer programs for multivariate data analysis in ecology and systematics. *Abstracta Botanica*, 17, p. 289-302.
- PODANI, János e FEOLI, Enrico **1991** - A general strategy for the simultaneous classification of variables and objects in ecological data tables. *Journal of Vegetation Science*, 2, p. 435-444.
- PRENTICE, I.C. **1977** - Non-metric ordination methods in ecology. *Journal of Ecology*, 65, p. 85-94.
- PROBST, JOHN R., RAKSTAD, DON S., RUGG, DAVID J. **1992** - Breeding bird communities in regenerating and mature broadleaf forests in the USA Lake States. *Forest Ecology and Management*, 49, p. 43-60.
- REYMENT, R. A., BLACKITH, R. E. , CAMPBELL, N. A. **1984** - Multivariate Morphometrics. Academic Press, London, 2ª Edition.
- RIBEIRO, ORLANDO e LAUTENSACH, HERMANN **1987** - Geografia de Portugal I. A Posição Geográfica e o Território. Edições João Sá da Costa, Lisboa.
- RIBEIRO, ORLANDO **1987** - Portugal o Mediterrâneo e o Atlântico. Livraria Sá da Costa Editora, Lisboa, 5ª Edição.

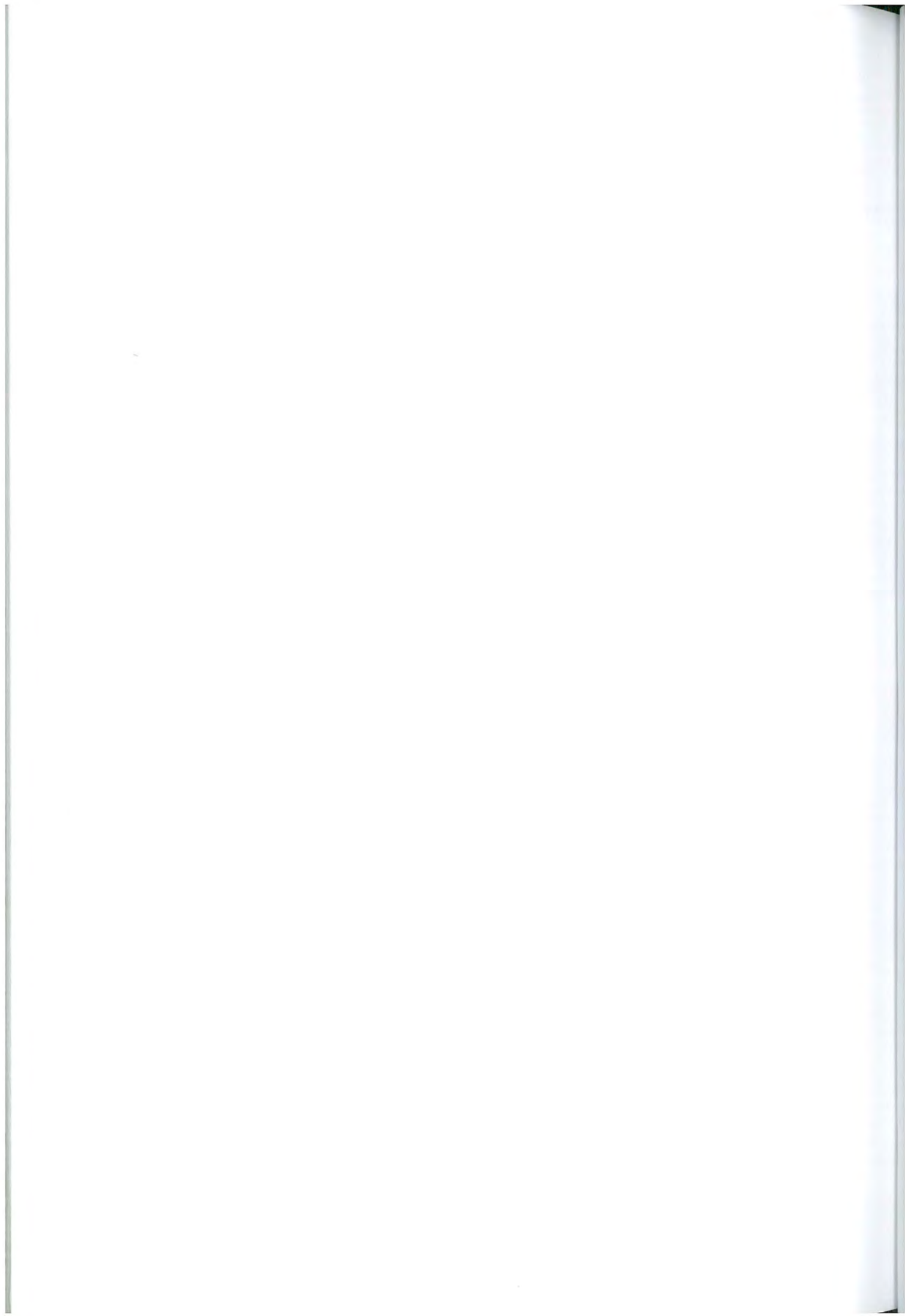
- ROUGERIE, G. e BEROUTCHACHVILI, N. **1991** - Géosystèmes et Paysages: Bilan et méthodes. Armand Colin Éditeur, Paris.
- SZARO, ROBERT C. **1990** - Southwestern riparian plant communities: Site characteristics, tree species distributions, and size-class structures. *Forest Ecology and Management*, 33/34, p. 315-334.
- SZARO, ROBERT C. e KING, RUDY M. **1990** - Sampling intensity and species richness: Effects on delineating southwestern riparian plant communities. *Forest Ecology and Management*, 33/34, p. 335-349.
- VINK, **1983** - Landscape Ecology and Land Use. Longman, London.
- VOS, W. e STORTELDER, A. **1992** - Vanishing Tuscan landscapes - Landscape ecology of a Submediterranean-Montane area (Solano Basin, Tuscany, Italy). Pudoc Scientific Publishers, Wageningen.
- WILLIAMS, W.T. e LAMBERT, J.M. **1961** - Multivariate methods in plant ecology. III. Inverse association-analysis. *Journal of Ecology*, 49, p. 717-729.
- WILLIAMS, W.T., LAMBERT, J.M., LANCE, G.N. **1966** - Multivariate methods in plant ecology. V. Similarity-analysis and Information-analysis. *Journal of Ecology*, 54, p. 427-445.
- ZONNEVELD, I. *in press* - Landscape Ecology (an introduction to Landscape Ecology as a base for Land Evaluation and Management)

Anexos



ANEXO I

- Localização aproximada dos meridianos e locais de amostragem.



ESPAÑA

Escala 1:750 000

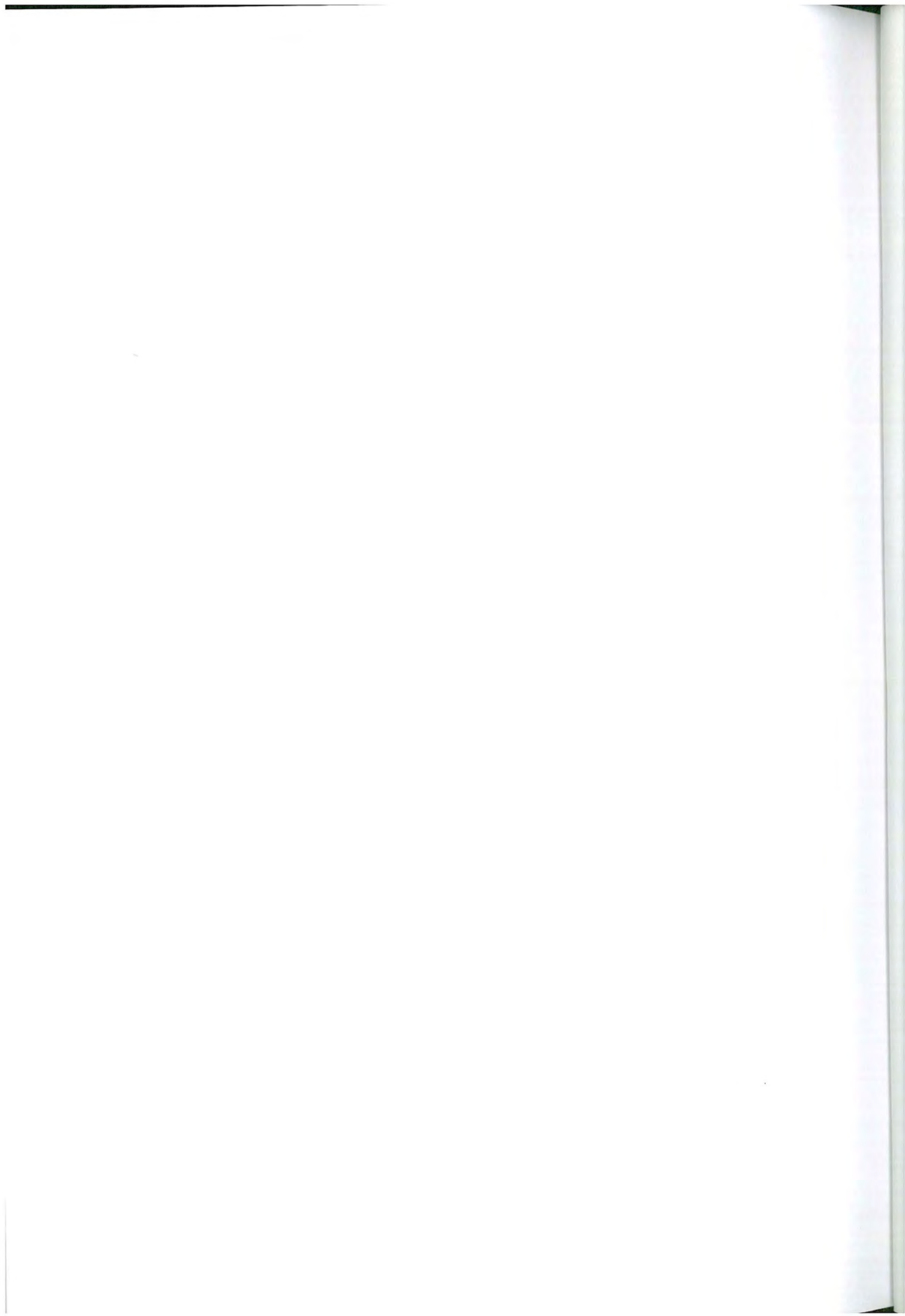
BEIRA ALTA

ENTRE
DOURO
E
MINHO



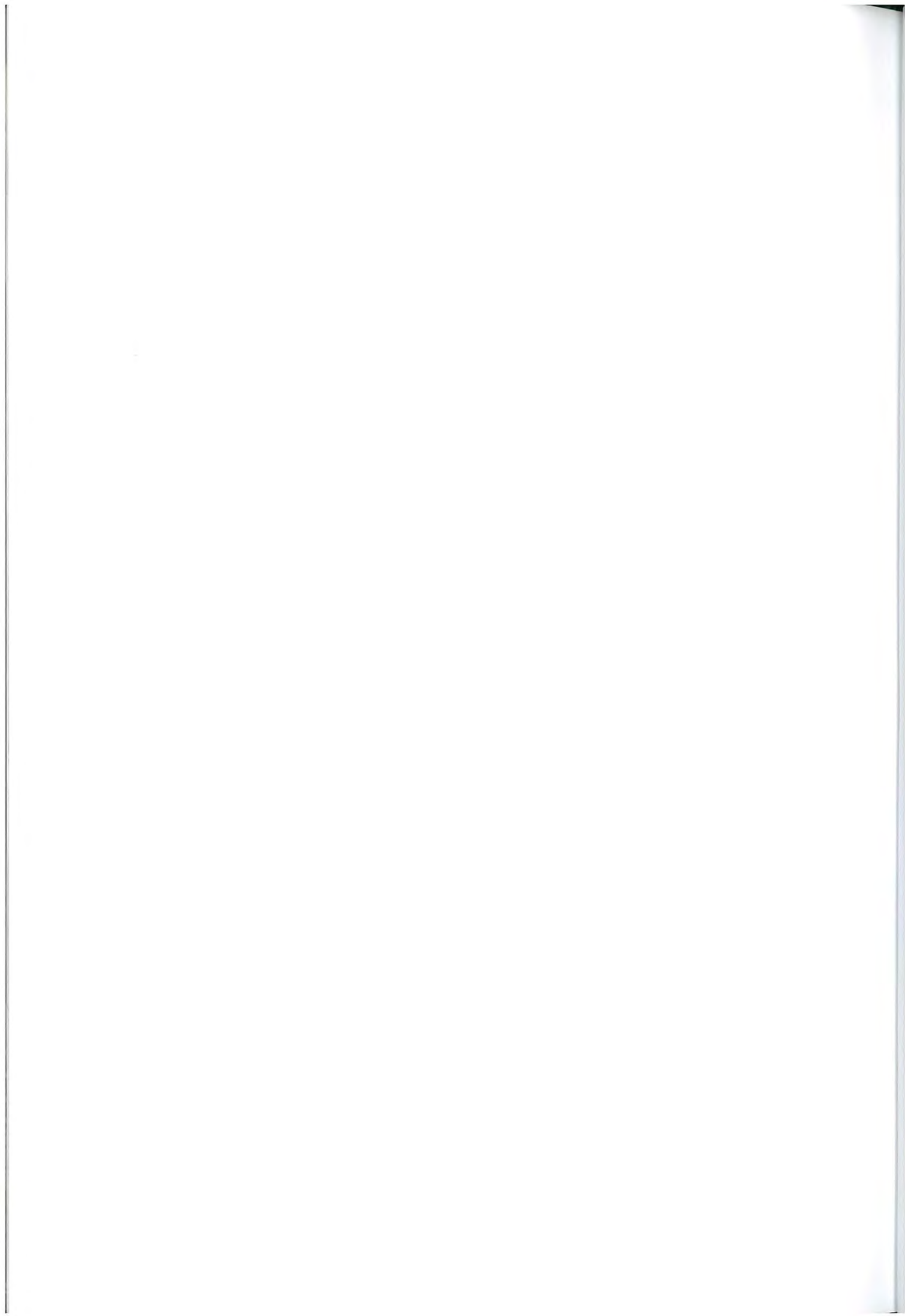
ANEXO 2

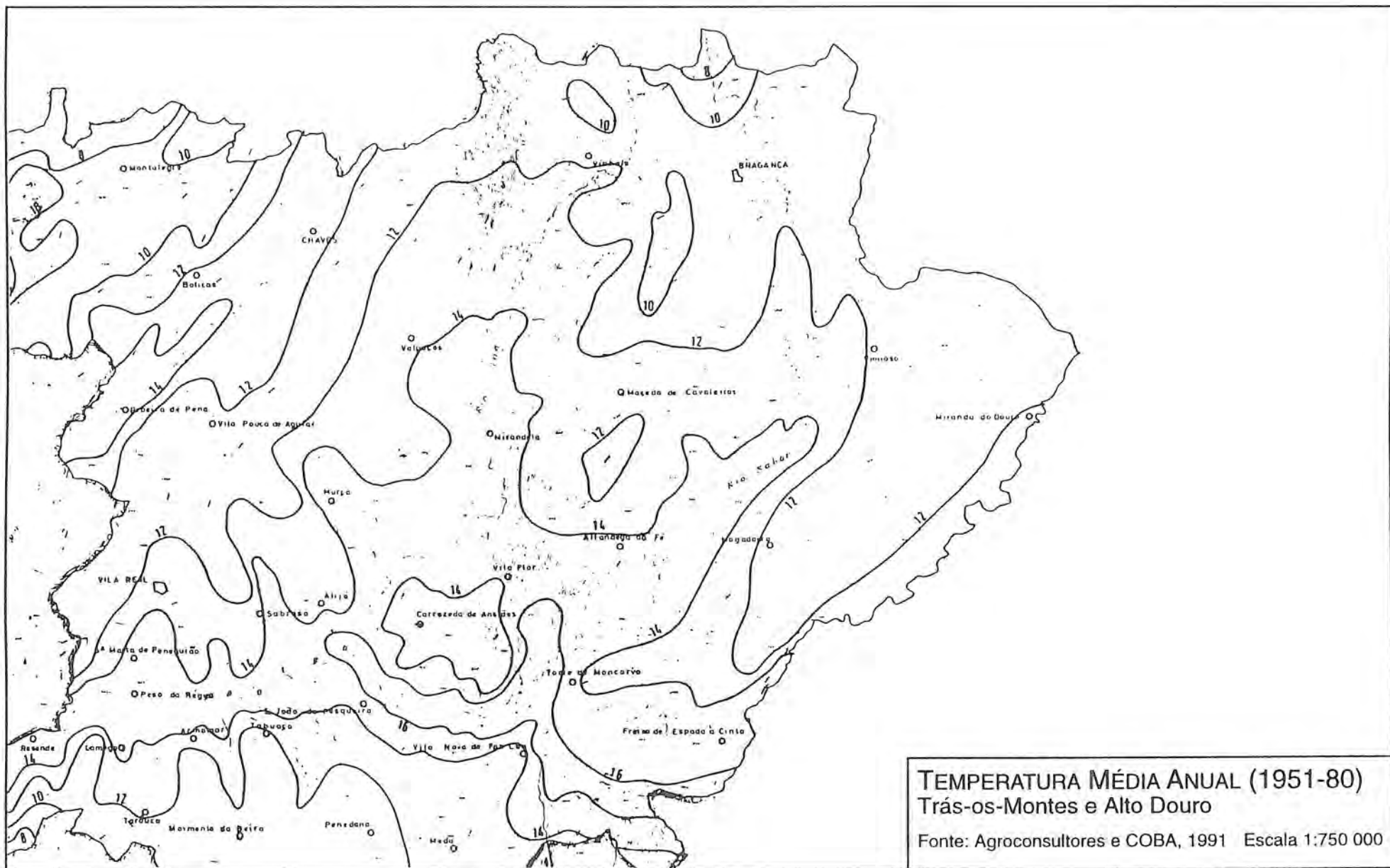
- *Tabela de Dupla Entrada resultante da análise de "TWINS PAN"*



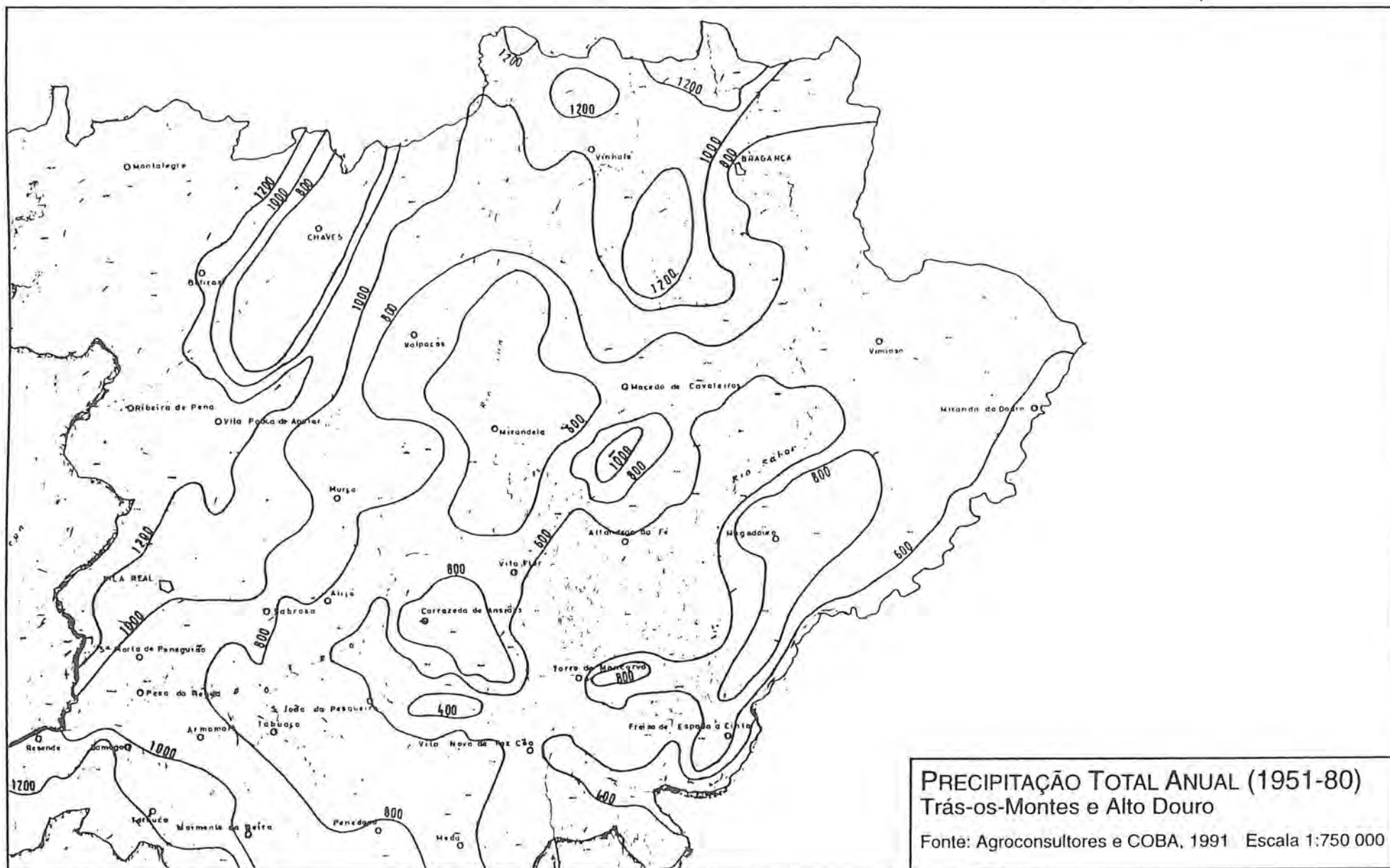
ANEXO 3

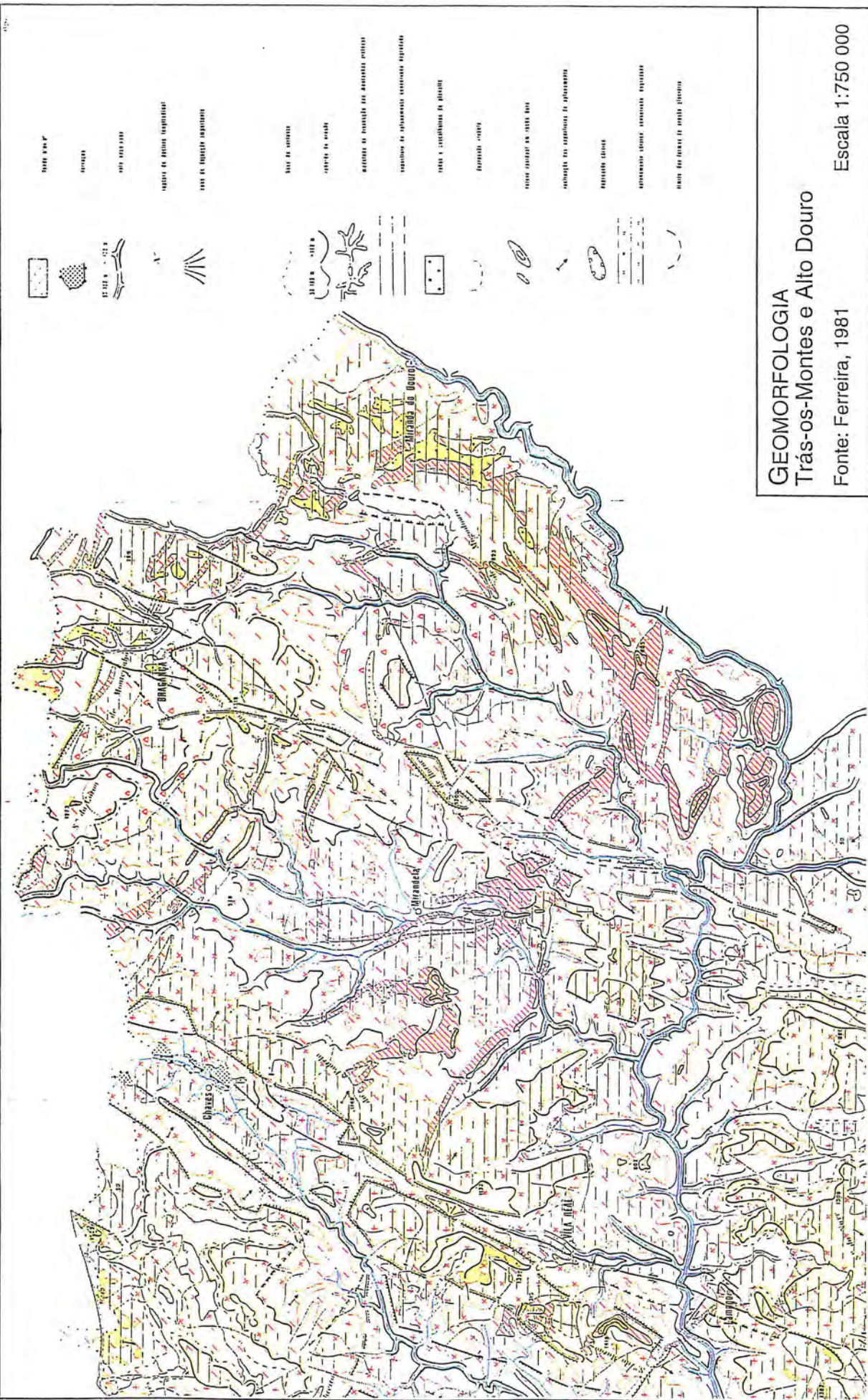
- *Mapa da Temperatura Média Anual*
- *Mapa da Precipitação Total Anual*
- *Mapa Geomorfológico*
- *Mapa Hipsométrico*
- *Mapa de Distribuição das Culturas Arvenses de Sequeiro*
- *Mapa de Distribuição de Carvalhos e Castanheiro*





TEMPERATURA MÉDIA ANUAL (1951-80)
Trás-os-Montes e Alto Douro
Fonte: Agroconsultores e COBA, 1991 Escala 1:750 000





ROCHA MATRIZ

CONCRETO

ROCHA ÍGNEA

ROCHA DE GRANITO (granito)

ROCHA DE GRANITO (granito)

ROCHA DE GRANITO (granito)

ROCHA DE GRANITO (granito)

ROCHA DE GRANITO (granito)

ROCHA DE GRANITO (granito)

ROCHA DE GRANITO (granito)

ROCHA DE GRANITO (granito)

ROCHA DE GRANITO (granito)

ROCHA DE GRANITO (granito)

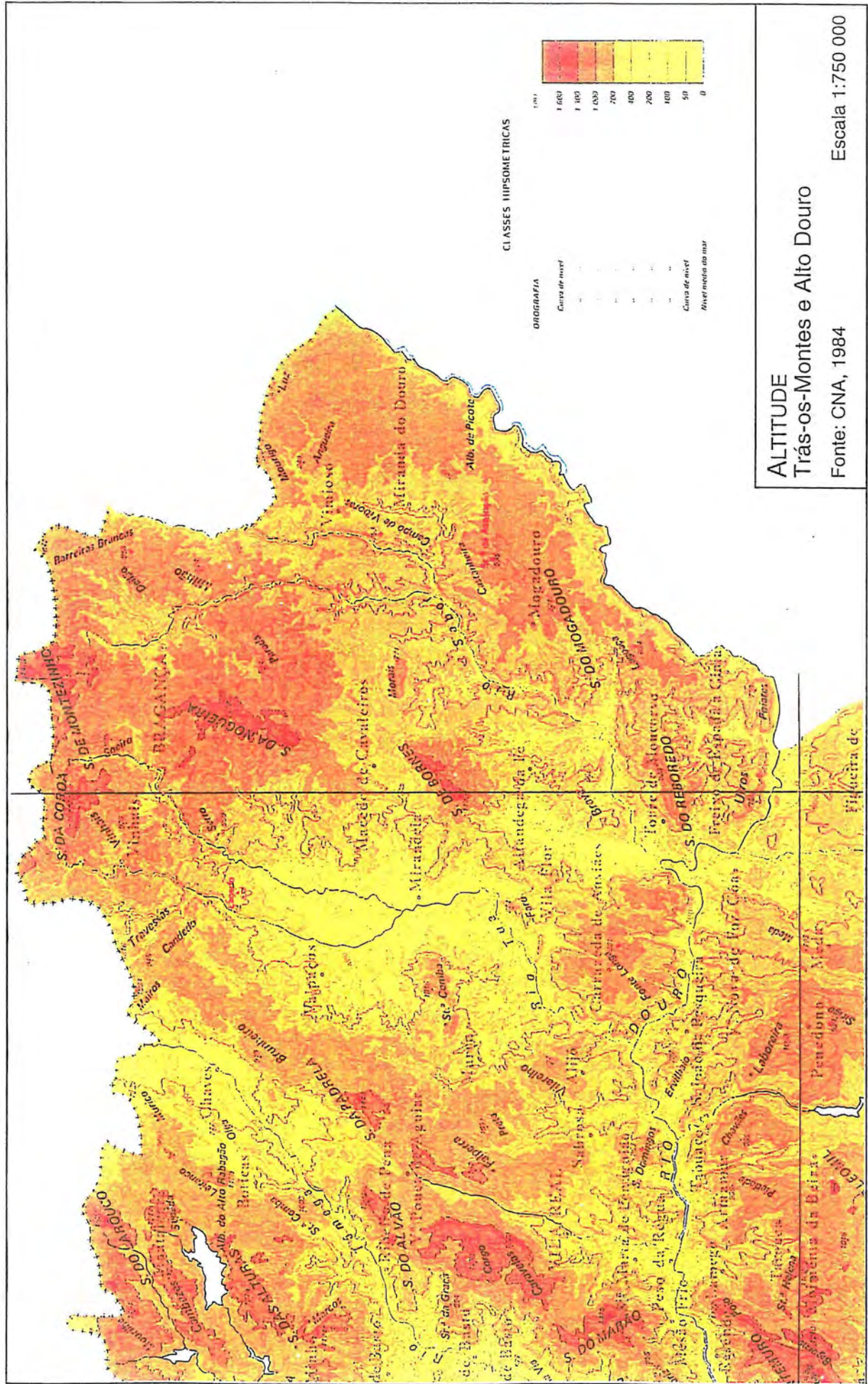
ROCHA DE GRANITO (granito)

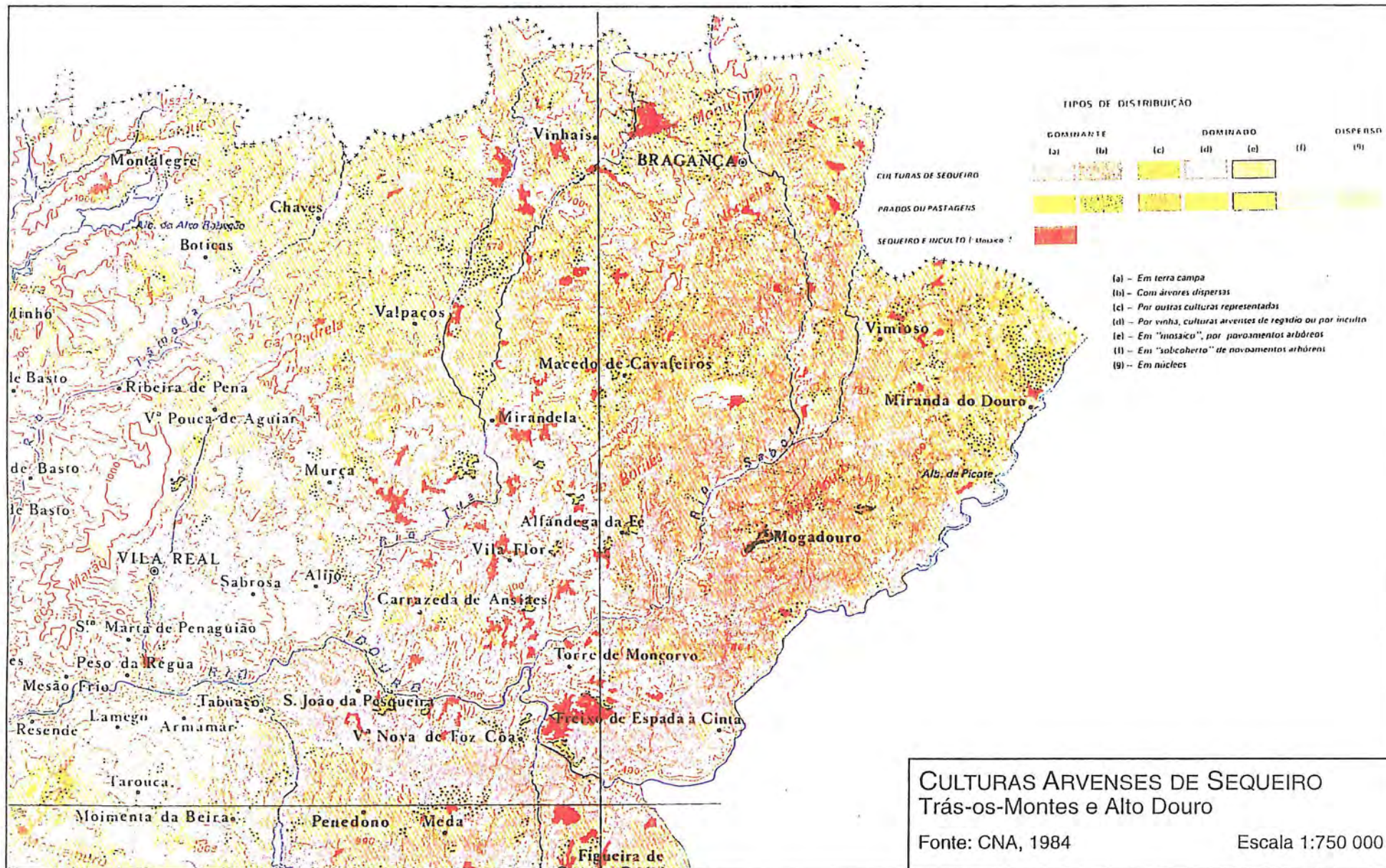
ROCHA DE GRANITO (granito)

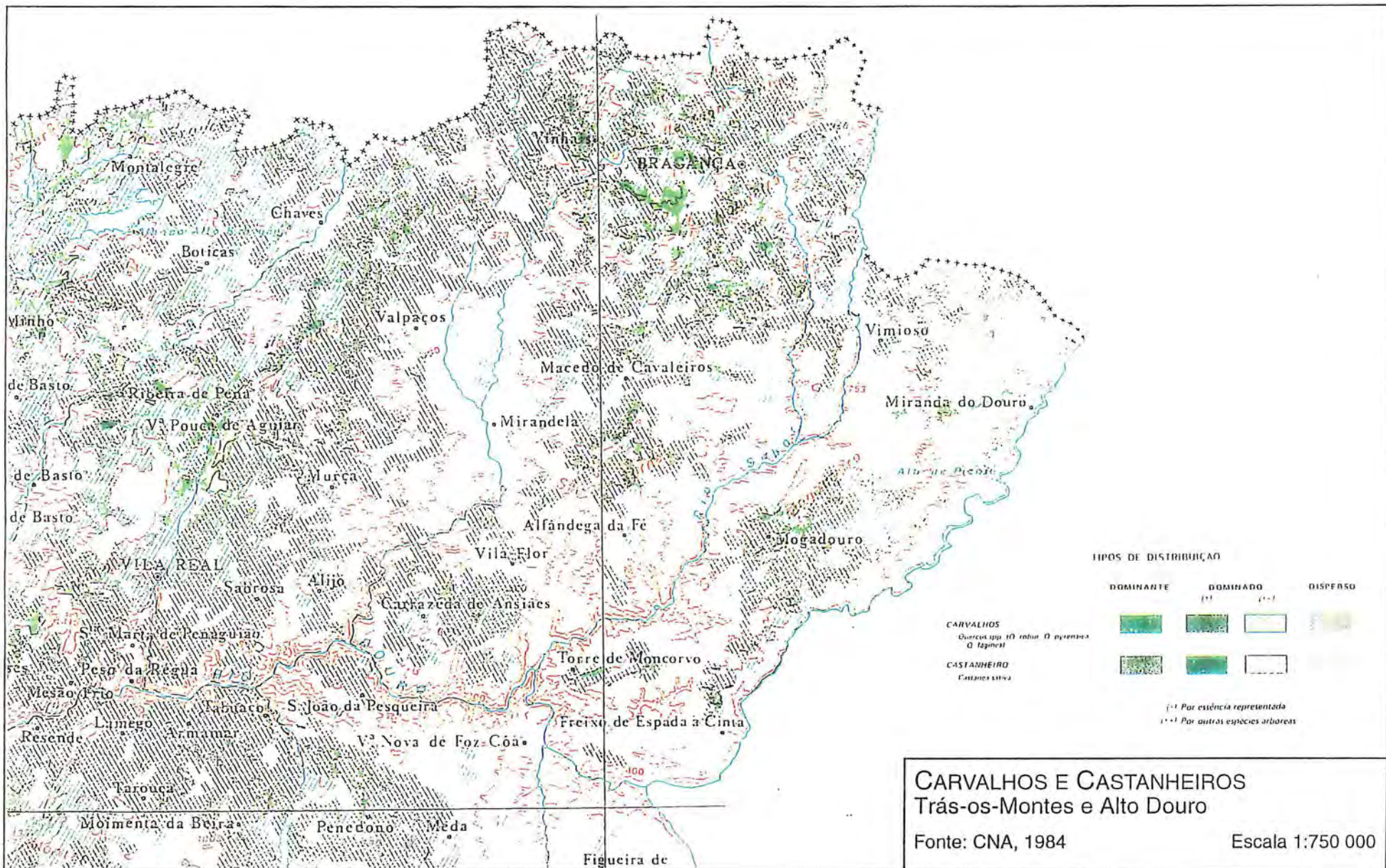
ROCHA DE GRANITO (granito)

GEOMORFOLOGIA
Trás-os-Montes e Alto Douro
 Fonte: Ferreira, 1981

Escala 1:750 000

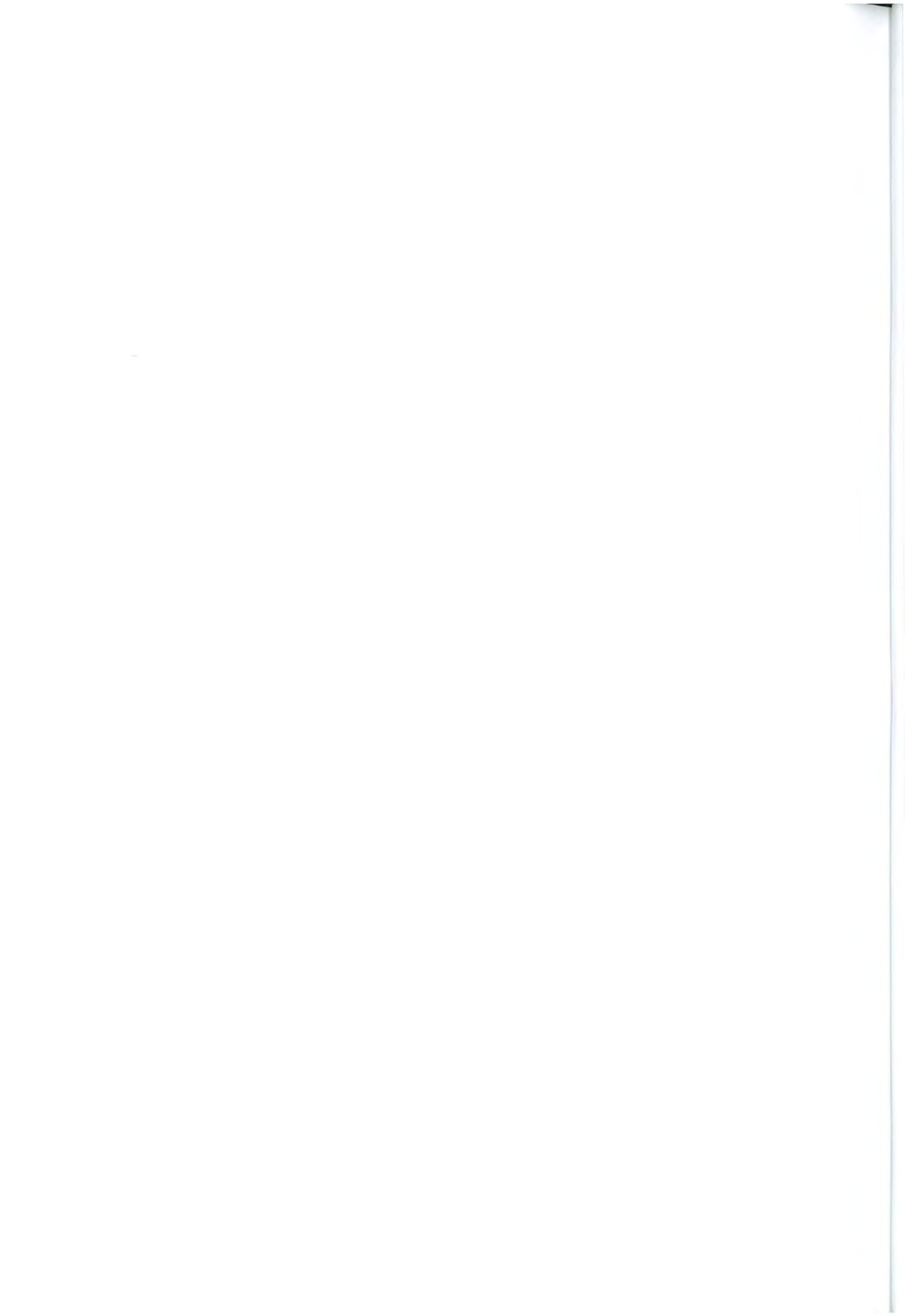


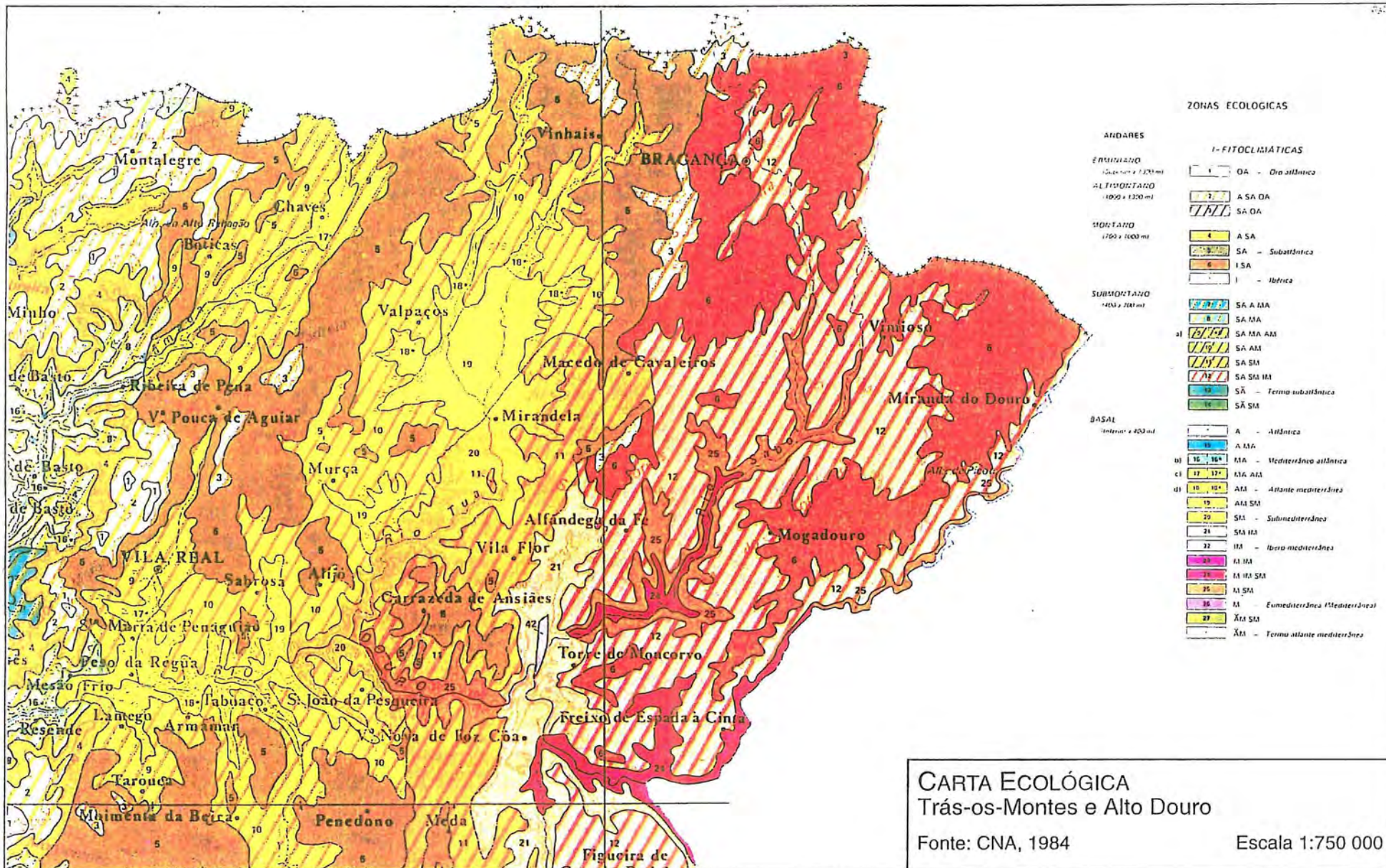




ANEXO 4

- *Classificação das Grandes Regiões*
- *Classificação da Carta Ecológica*
- *Classificação da Carta de Agrotipos*





CARTA ECOLÓGICA
Trás-os-Montes e Alto Douro

Fonte: CNA, 1984

Escala 1:750 000

