

Identificação dos Distritos Industriais em Portugal

João Cerejeira¹²

Resumo

Vários autores têm procurado a definição de uma metodologia de identificação dos sistemas locais de produção, nomeadamente em termos da operacionalização da noção de "distrito industrial". Dentro dos trabalhos mais citados na literatura é de destacar o contributo pioneiro de Sforzi (1987 e 1990), onde identifica 61 distritos industriais em Itália, recorrendo à análise estatística multivariada.

No caso português, para além dos estudos de caso, não são conhecidos trabalhos de classificação sistemática, à escala do concelho, de forma a tipificar e identificar os sistemas produtivos locais a nível nacional.

Assim, neste artigo, propomos a utilização, como algoritmo de classificação, a análise de *clusters* numa variante apropriada à análise espacial, efectuando a segmentação de todo o país em grupos de concelhos mais ou menos homogêneos, relativamente a determinadas características pré-determinadas, de forma a identificar aqueles que possam ser classificados como distritos industriais.

1. Introdução

Vários autores tem procurado a definição de uma metodologia de identificação dos sistemas locais de produção, nomeadamente em termos de operacionalização da noção de "distrito industrial". Dentro dos trabalhos mais citados na literatura são de destacar os contributos pioneiros de Garofoli (1983 e 1994) e de Fabio Sforzi (1987 e 1990). No primeiro caso, trata-se de procurar zonas de especialização produtiva e de pequenas empresas, a uma escala equivalente ao concelho, de forma a serem classificadas segundo a sua natureza mais ou menos sistémica, em três categorias: áreas de especialização produtiva, sistemas produtivos locais e áreas-sistema. No segundo caso, o objectivo principal é a identificação dos distritos industriais Marshallianos existentes em Itália, a partir do conjunto das 955 *Local Labour Market Areas* (LLMAs), unidades espaço-funcionais próximas do conceito de "bacia de emprego", definidas através da regionalização funcional dos fluxos casa-trabalho. Sforzi (1987) utiliza a análise multivariada, procedendo à classificação das LLMA em 15 categorias, sendo numa desta onde se identificam 61 distritos industriais, quase todos localizados no Nordeste e Centro da Itália.

No caso português, para além de vários estudos de caso, não são conhecidos trabalhos de classificação sistemática, à escala do concelho, de forma a tipificar os sistemas produtivos locais, ou, se preferirmos, os distritos industriais, a nível nacional. Refira-se, no entanto, a análise da configuração da região Norte Litoral (48 concelhos), presente em Silva (1988) e Silva e Figueiredo (1992), onde se identifica um articulado de espaços produtivos locais de diferentes níveis de industrialização e complexidade.

Assim, neste trabalho, propomo-nos efectuar a segmentação de todo o país em grupos de concelhos, mais ou menos homogêneos, relativamente a determinadas características pré-definidas, de forma a identificar aqueles que possam ser classificados como distritos industriais. Em termos metodológicos,

¹ Escola de Economia e Gestão - Universidade do Minho
e-mail: jccsilva@eeg.uminho.pt

² O autor agradece o apoio da *Fundação para a Ciência e Tecnologia*, através do projecto P/ECO/13134/1998, no âmbito do Programa PRAXIS.

serão utilizados procedimentos da estatística multivariada e, como técnica de classificação, a análise de *clusters*, numa variante apropriada para a análise espacial. É de notar que a utilização de métodos de estatística multivariada para classificação de unidades espaciais tem sido bastante reduzida em Portugal apesar de existirem vários trabalhos aplicados a outros países. Na verdade, apenas se conhecem, no caso português, os trabalhos de Brandão, Pires e Portugal (1998), e, para o caso particular da Região Norte, de Gomes, Bacelar e Saleiro (1994), ambos utilizando informação do Recenseamento da População de 1991, e com objectivos meramente descritivos das regiões resultantes, em termos sociais, económicos e políticos.

2. A análise de *clusters* como método de classificação na análise espacial

A importância da classificação parece evidente em ramos do conhecimento onde a categorização de elementos é uma base estrutural, como são o caso da biologia (taxonomia das espécies, p.ex.⁰) ou da medicina (classificação de doenças, p.ex.⁰). Genericamente, designa-se por análise de *clusters* um conjunto de métodos de estatística multivariada que permitem agrupar uma população de n indivíduos, caracterizados por q variáveis, em $k < n$ conjuntos relativamente homogéneos, designados genericamente por *clusters*, observando apenas as semelhanças ou dessemelhanças entre eles³. Nas ciências sociais são conhecidos trabalhos de aplicação da análise de *clusters* em áreas tão diversas como a antropologia, a ciência política, ou as ciências empresariais⁴.

No âmbito da análise espacial, existem inúmeras situações em que é útil a classificação, ou a tipificação, das unidades espaciais de base em categorias, de acordo com determinadas características pré-determinadas. Beguin (1979) refere cinco objectivos passíveis de um processo de classificação de unidades espaciais:

- (1) o objectivo da simples redução do número de unidades a analisar;
- (2) o objectivo de definição de regiões homogéneas, ou seja de conjuntos de locais contíguos com características semelhantes;
- (3) o objectivo de definição de uma tipologia das unidades espaciais, onde não se impõe uma restrição de vizinhança ou contiguidade como no caso precedente;
- (4) o objectivo em termos de análise exploratória de forma a gerar hipóteses de pesquisa;
- (5) e, por último, o objectivo de testar uma hipótese previamente formulada.

Como é natural, diferentes critérios de categorização implicam diferentes resultados. Cliff *et al.* (1975) sugerem, em geral, que uma categorização "óptima" de unidades geográficas, deverá simultaneamente satisfazer os seguintes critérios:

- a) Deverá ser simples, no sentido de encontrar uma solução que produza o mínimo de categorias;
- b) As unidades espaciais classificadas na mesma categoria deverão ser relativamente semelhantes, em termos das variáveis que as caracterizam;

³ A contribuição mais expressiva para a aplicação destes métodos foi dada por Robert Sokal e Peter Sneath em 1963 com a obra *Principles of Numerical Taxonomy*, orientada precisamente para a classificação biológica.

⁴ Como exemplos de aplicações da análise de *clusters* nas ciências empresariais temos a segmentação de mercados ou a identificação de grupos de empresas com comportamentos estratégicos semelhantes (ver Reis, 1997).

c) Se o objectivo for a constituição de regiões, deverá ser respeitado o princípio da contiguidade entre unidades classificadas na mesma categoria.

2.1. A análise de clusters

De modo sintético, o método de análise de *clusters* procede ao agrupamento dos objectos em função da informação existente, de tal modo que os objectos pertencentes a um mesmo grupo sejam tão semelhantes quanto possível, e sempre mais semelhantes aos elementos do mesmo grupo do que a elementos dos restantes grupos.

Apesar de não existir um único critério de partição e/ou agrupamento dos casos, de existirem diferentes objectivos nas análises, ou diferentes tipos de dados iniciais, em todas as aplicações da análise de *clusters* são cumpridas cinco etapas principais:

- (1) a selecção da amostra dos objectos a serem agrupados;
- (2) a definição de um conjunto de variáveis que caracterizem os objectos da amostra;
- (3) a escolha de uma medida de semelhança ou distância entre cada par de objectos;
- (4) a escolha e utilização de um algoritmo de classificação;
- (5) e, por último, a validação dos resultados encontrados.

De seguida, discutem-se os aspectos fundamentais de cada uma das etapas acima descritas.

2.2. A selecção da amostra e das variáveis caracterizadoras do objecto da amostra

A selecção da amostra dos objectos a agrupar e das variáveis que os caracterizam conduzem a um duplo problema: por um lado, um problema substantivo, que resulta da necessidade de escolher por entre os dados disponíveis, aqueles que serão utilizados no método de classificação, problema que será resolvido através do conhecimento prévio do investigador acerca do objecto de estudo; por outro lado, existe um problema de cariz estatístico que tem a ver com a possível utilização de variáveis definidas em escalas de medida diferentes. Apesar de existir alguma controvérsia relativa aos procedimentos a utilizar perante uma situação deste tipo⁵, caso se aceite uma distribuição normal, e que as variáveis são independentes entre si, então o processo mais utilizado consiste na standardização das variáveis, através da sua transformação em novas variáveis do tipo $Z = (X - m) / s$ em que X representa a variável original, m representa a média amostral da variável e s o seu desvio padrão. É de notar que este processo de standardização, ao impor uma média nula e um desvio padrão unitário para todas as variáveis transformadas, reduz as diferenças entre os indivíduos, ficando todas as variáveis com o mesmo peso.

Na hipótese de existirem variáveis com uma importância intrínseca superior, é possível proceder a sua ponderação, que pode justificar-se no caso de existirem argumentos teóricos suficientemente fortes para tal.

⁵ Ver Aldenderfer e Blashfield (1984, p 20).

2.3. A escolha de uma medida de semelhança ou distância entre cada par de objectos

As medidas de (dis)semelhança habitualmente utilizadas nas ciências sociais são as medidas de distância ou dissemelhança entre os elementos de uma matriz de dados. Dentro destas, e apesar de existirem várias métricas possíveis, é normalmente utilizada como medida de distância entre dois casos i e j , a *Distância Euclideana*: definida como a raiz quadrada do somatório dos quadrados das diferenças entre os valores de i e de j para todas as variáveis ($v = 1, 2, \dots, j$),

$$d_{ij} = \sqrt{\sum_{v=1}^q (x_{vi} - x_{vj})^2}. \quad (1)$$

A utilização de qualquer métrica obriga à standardização das variáveis, a qual, como vimos, não é isenta de alguns problemas. De notar, também, que os resultados obtidos por um qualquer algoritmo de classificação estarão obrigatoriamente dependentes da métrica utilizada.

2.4. Identificação de clusters na análise espacial

Apesar de não existir uma designação exacta para o termo *cluster*, é consensual a proposta de Everitt (1980), na qual *clusters* são "... continuous regions of [a] space containing a relatively high density of points, separated from other such regions by regions containing a relatively low density of points." Assim, um problema comum a qualquer análise é a escolha de um critério de (des)agregação dos casos com o objectivo de efectuar a formação de grupos cujos elementos sejam similares entre si.

Existem várias "famílias" de técnicas de *clustering*, cada uma representando uma perspectiva diferente de formação dos grupos, que se reflecte nos diferentes algoritmos de classificação que são utilizados. Desta forma, é frequente que os resultados obtidos não sejam idênticos quando se aplicam diferentes métodos sobre uma mesma amostra. Apesar de não existir aquilo a que se possa chamar o melhor método de *clustering*, uma vez que todos têm vantagens e desvantagens, determinados tipos de técnicas são particularmente usados na resolução de problemas específicos.

No caso da análise espacial, os trabalhos mais recentes apontam para a utilização de métodos de optimização (Wise, Haining e Ma, 1997; Murray, 1998), em detrimento das técnicas hierárquicas, sendo o mais conhecido o método *k-means* (MacQueen, 1967).

Existem três modelos de optimização que poderão ser utilizados em problemas de classificação espacial⁶. O mais conhecido e utilizado é o método CPCP⁷, uma variante para análise espacial do método *k-means* (MacQueen, 1967), baseado no trabalho de Cooper (1963), e que poderá ser encontrado na maioria dos *packages* estatísticos. Este método consiste no agrupamento das observações espaciais baseado na minimização da sua distância Euclideana a pontos centrais criados artificialmente. Podemos, assim, usar a seguinte notação:

i = índice de localização de cada uma das n observações ($i = 1, \dots, n$), caracterizadas por um vector \mathbf{V} de variáveis;

p_p = ponto central (centróide) do *cluster* p ($p = 1, \dots, k$);

d_{ip} = distância Euclideana entre a observação localizada em i e p_p .

⁶ Para uma revisão recente destes modelos ver Murray e Estivill-Castro (1998).

⁷ CPCP – *Center Points Clustering Problem*.

$$y_{ip} = \begin{cases} 1 & \text{se a observação localizada em } i \text{ pertence ao cluster } p \\ 0 & \text{caso contrário} \end{cases}$$

O objectivo do CPCP é minimizar a diferença total entre as coordenada de cada observação (valores observados da variáveis) e as coordenadas dos pontos centrais:

$$\min Z = \sum_{i=1}^n \sum_{p=1}^k d_{ip} y_{ip}, \quad (2)$$

com a restrição de que todas as observações ficam afectas a um e um só *cluster*:

$$\sum_{p=1}^k y_{ip} = 1. \quad (3)$$

A heurística mais conhecida e aplicada, no caso do método CPCP, é a seguinte (Murray, 1998):

1º passo: Escolher uma partição inicial das unidades espaciais por um número k de *clusters* indicado pelo investigador;

2º passo: Calcular os respectivos centróides;

3º passo: Calcular a distâncias entre cada observação e os centróides dos vários grupos, transferindo cada indivíduo para o grupo relativamente ao qual se encontra a uma menor distância;

4º passo: Se no passo anterior não foi efectuada nenhuma transferência de observações de um *cluster* para outro, a heurística termina, tendo sido encontrado um óptimo local. No caso contrário, voltar ao 2º passo.

Uma das maiores limitações destas técnicas, é que a sua *performance* depende da escolha de uma partição inicial, de forma a evitar uma solução sub-óptima⁸. Porém, Milligan (1998) demonstrou que o método *k-means*, quando utiliza uma partição inicial derivada da aplicação de uma técnica hierárquica, obtém uma *performance* superior aos métodos hierárquicos, em termos da obtenção de um óptimo global e não um óptimo local.

2.5. Validação dos resultados

Uma vez que a análise de *clusters* tem como objectivo criar grupos homogéneos, surge este problema, que é o da escolha do número de *clusters*, no caso de não existir o conhecimento do número de grupos em que a população em estudo se deverá dividir. Um dos métodos usados quando se usam técnicas hierárquicas, consiste na comparação gráfica do número de *clusters* com o respectivo coeficiente de fusão, isto é, o valor numérico (distância ou semelhança) para o qual vários casos se unem para formar um grupo. Assim, quando a divisão de um novo grupo não introduz alterações significativas no coeficiente de fusão, poderá tomar-se essa partição como sendo óptima.

Outro procedimento utilizado é o da comparação dos resultados obtidos por vários critérios diferentes de agrupamento. Poder-se-á aferir o grau de convergência entre os vários critérios de agrupamento através de uma tabela de

⁸ A possibilidade da solução ser um óptimo local, e não global, resulta da incapacidade prática em gerar todas as possíveis combinações de partição, das n observações em k grupos, de forma a escolher aquela que minimiza o somatório das distâncias entre as observações e os centróides de cada grupo. Assim, a generalidade dos métodos de optimização são heurísticos, havendo várias alternativas para a heurística a utilizar.

contingência, indicando o número de observações que se agrupam no mesmo *cluster*, para o mesmo número de *clusters*. Desta forma é possível verificar a maior ou menor estabilidade das soluções encontradas, de maneira a concluir acerca da qualidade do agrupamento efectuado.

3. Aplicação da análise de *clusters* para identificação dos distritos industriais em Portugal

A identificação dos distritos industriais em Portugal, seguirá quatro etapas principais. Em primeiro lugar há que optar por um conceito operacional de distrito industrial, ou seja, um conceito passível de ser caracterizado por variáveis quantitativas. Uma segunda etapa consiste na escolha das variáveis, numa dada escala de análise, e sua quantificação. Um terceiro passo será a utilização de uma técnica de classificação e posterior escolha do(s) grupo(s) cujas características mais se assemelham ao conceito de distrito industrial proposto. A quarta e última etapa será a validação dos resultados obtidos, através da identificação dos distritos industriais presentes nestes grupos, a partir do seu perfil de especialização, e do recurso a outras variáveis de caracterização complementares.

3.1. O conceito de distrito industrial

Quanto à definição operacional de distrito industrial, é de recordar o conceito avançado por Bianchi (1998), onde é proposta uma definição baseada numa sobreposição de três modelos interdependentes: o modelo de produção, no qual se destaca a característica da flexibilidade, em sentido lato, induzida por uma organização produtiva dominada por pequenas empresas especializadas numa fase ou componente específica do produto, operando numa mesma fileira produtiva. A esta especialização, está associada uma concentração espacial dos vários actores económicos (modelo de aglomeração espacial), possibilitando uma redução dos custos de transacção e de reprodução e qualificação da mão-de-obra. Forma-se, assim, uma rede densa e complexa de relações entre os agentes, assente numa forte homogeneidade cultural e num consenso social, que permite a regulação local das relações de produção (modelo social).

3.2. As variáveis

Para proceder à identificação dos distritos industriais em Portugal, é, em segundo lugar, necessário criar uma tipologia das unidades espaciais (no nosso caso, os 275 concelhos do continente), utilizando um conjunto de variáveis que consigam captar as características fundamentais do modelo de distrito industrial. No nosso caso, a escolha incidiu sobre quatro variáveis que permitem evidenciar quatro destas características, nomeadamente o perfil industrial e a especialização dos concelhos, a existência de uma concentração de trabalhadores em pequenas e médias empresas, espacialmente concentradas, e por último, a aglomeração geográfica de estabelecimentos do mesmo ramo industrial. Notemos que estas variáveis caracterizam essencialmente o modelo de produção e de aglomeração espacial, referidos no ponto anterior. Quanto às variáveis características do modelo social, serão referidas na última etapa, ou seja como auxiliares na validação dos resultados obtidos.

As variáveis escolhidas, resumidas na tabela seguinte, tem como unidade de análise o concelho (unidade geográfica mais pequena para a qual foi possível obter os dados estatísticos necessários), e são as seguintes:

- a *Taxa de Industrialização* (PEMPIT), medida como a proporção do emprego na indústria transformadora relativamente à população empregada do concelho;
- o *Coefficiente de Especialização Industrial* (COEFESP), que é um indicador que compara a estrutura industrial de cada concelho com a estrutura de um agregado de referência (no caso o continente);
- a *Densidade de Emprego em Pequenas e Médias Empresas* (PME50KM), medida pelo número de trabalhadores em estabelecimentos industriais com menos de 50 empregados, por Km²;
- a *Especialização e Aglomeração Industrial* (ESPINDKM), calculada pelo número de estabelecimentos do ramo mais importante do concelho, por Km²;

Uma vez que as variáveis estão em escalas diferentes, procedeu-se à sua standardização prévia, através da transformação já anteriormente enunciada.

Tabela 1: Descrição das Variáveis

Variável	Descrição	Data	Fonte
PEMPIT	Taxa de Industrialização	$\frac{X_i}{PE_i}$ Média 1990_92 1991	Q.P. I.N.E.
COEFESP	Coefficiente de Especialização Industrial	$\sum_j \left(\frac{X_{ij}}{X_i} - \frac{X_j}{X} \right)^2$ Média 1990_92	Q.P.
PME50KM	Densidade de Emprego em P.M.E.'s	$\frac{\bar{X}_i}{A_i}$ Média 1990_92	Q.P. I.N.E.
ESPINDKM	Especialização e Aglom. Industrial	$\frac{Emax_i}{A_i}$ Média 1990_92	Q.P. I.N.E.

Notas:

X_{ij} = Emprego no sector industrial j , no concelho i .
 X_i = Emprego total na indústria transformadora, no concelho i .
 PE_i = População empregada do concelho i .
 \bar{X}_j = Emprego no sector industrial j , no continente.
 \bar{X}_i = Emprego em estabelecimentos com menos de 50 trabalhadores, no concelho i .
 $Emax_i$ = N° de estabelecimentos do sector industrial com maior número, do concelho i .
 A_i = Área do concelho i .
 Q.P. = Quadros de Pessoal do Ministério do Trabalho e da Solidariedade.

3.3. Classificação usando a análise de clusters e resultados obtidos

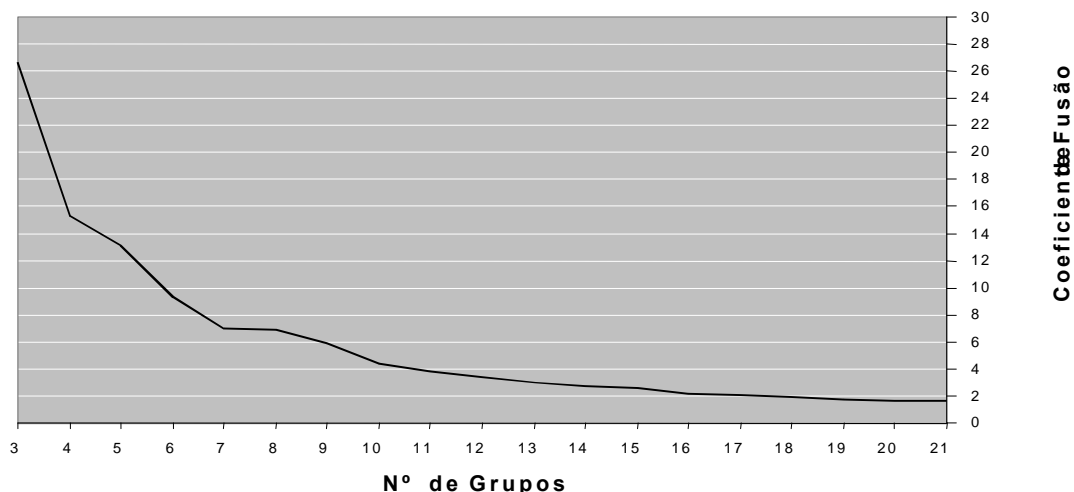
Tal como sugerido anteriormente, utilizou-se como método de classificação o método CPCP, na sua variante combinatória e inclusiva, ou seja, o centróide dos *clusters* é recalculado sempre que há uma alteração na sua composição, incluindo (ou excluindo) no seu cálculo as coordenadas da observação que foi incluída (ou excluída) no *cluster* a que se refere.

A escolha do número de grupos foi efectuada através de uma classificação prévia por um método hierárquico, segundo o critério da média inter-grupos (ver Reis, 1997, p. 317). Pela observação directa do gráfico 1^o que conjuga o número de *clusters* com o respectivo coeficiente de fusão, aceita-se como satisfatória a

⁹ Para permitir uma melhor visualização do gráfico, apresentamos apenas a secção mais relevante.

hipótese de partição dos concelhos em sete grupos (coeficiente de fusão igual a 6,94), tendo sido extraídos os respectivos centróides para utilização como centros iniciais na implementação do método de optimização.

Gráfico 1: Análise Gráfica do Coeficiente de Fusão



Na tabela 2 apresentam-se os centros dos *clusters* finais, bem como o número de elementos de cada grupo. A distribuição dos concelhos pelos vários *clusters* poderá ser observada no mapa da figura seguinte, o qual permite ter uma visão global e mais esclarecedora dos resultados obtidos.

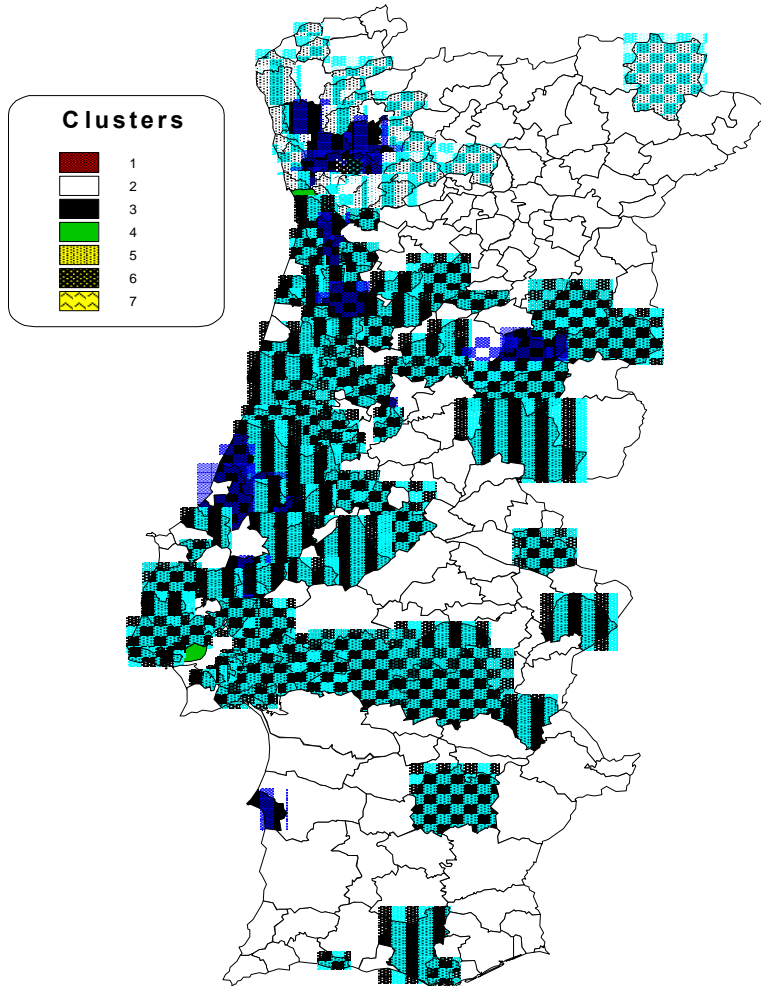
Tabela 2: Centros dos Clusters Finais

Variável (estandardizada)	Cluster							Mediana
	1	2	3	4	5	6	7	
Z(PEMPIT)	2,30	-0,51	7,08	0,49	0,19	2,06	1,29	-0,36
Z(COEFESP)	0,15	0,51	-0,13	-1,25	-0,79	0,01	0,82	-0,25
Z(PME50KM)	0,27	-0,22	11,80	6,70	0,00	7,86	0,82	-0,22
Z(ESPINDKM)	0,31	-0,19	13,94	3,33	-0,06	5,35	2,72	-0,20
Nº de elementos	20	151	1	2	99	1	1	

Utilizando como medida de comparação a mediana de cada variável, por ser uma medida de localização mais robusta que a média, podemos genericamente caracterizar os grupos da seguinte forma:

(1) *Clusters* 1, 3, 6 e 7 - Neste conjunto, os 4 *clusters* possuem centróides que atingem valores mais elevados que a mediana, em todas as variáveis consideradas. Os 23 concelhos que compõem estes *clusters* localizam-se em quatro zonas distintas: na parte Norte de um anel em torno da área metropolitana do Porto (Vale do Ave, parte do Vale do Cávado - concelho de Barcelos - e Vale do Sousa); na parte Sul desse mesmo anel (zona de Entre-Douro-e-Vouga); na região envolvente à Serra da Estrela; e, por último, na zona Norte da Estremadura. Genericamente podemos afirmar que estes *clusters* se caracterizam por uma elevada taxa de industrialização, associada a uma população de pequenas e médias empresas espacialmente aglomeradas, possuindo também uma estrutura industrial bastante especializada e uma elevada densidade de estabelecimentos industriais do mesmo ramo por Km². Assim, podemos constatar ser este o grupo de *clusters* cujas características mais se assemelham ao modelo de distrito industrial, ou seja, é neste conjunto de concelhos que irão ser identificados os núcleos dos distritos industriais portugueses.

Figura 1: Distribuição dos Concelhos do Continente pelos Vários Clusters



(2) *Cluster 2* - É o grupo com um maior número de elementos (151). É constituído por concelhos com uma baixa taxa de industrialização e uma baixa concentração de unidades industriais. Sendo concelhos com um reduzido número de estabelecimentos, é natural a ocorrência de valores elevados para o Coeficiente de Especialização, uma vez que este indicador refere-se à especialização intra-industrial. Em termos geográficos, os concelhos deste *cluster* localizam-se predominantemente no Interior e Sul do País, ou seja em áreas onde os sectores primário e terciário são dominantes.

(3) *Cluster 4* - Este *cluster* é constituído pelos concelhos de Lisboa e Porto São caracterizados por uma concentração elevada de estabelecimentos industriais e por uma presença significativa de pequenas e médias empresas. A estrutura industrial existente é diversificada e a sua taxa de industrialização, apesar de superior à média nacional, é inferior à média verificada nos *clusters* 1, 3, 6 e 7.

(4) *Cluster 5* - é o segundo maior agrupamento de concelhos encontrado. Fazem parte deste grupo concelhos localizados no Litoral a Norte de Setúbal, e os concelhos das capitais de distrito, ou próximos destes, do Interior e Sul do País. É o *cluster* cujas médias das variáveis mais se aproximam das médias nacionais. São zonas com uma estrutura industrial diversificada (embora menor que no *cluster* anterior), e um nível de industrialização intermédio.

3.4. Validação dos resultados obtidos

Após a escolha dos *clusters* cujas características melhor se ajustam à definição proposta de distrito industrial, importa agora seleccionar quais destes concelhos poderão ser identificados como "núcleos" dos distritos industriais. Procedemos a esta selecção, excluindo ao grupo de 23 concelhos que constituem os *clusters* 1, 3, 6, e 7, os concelhos que apresentem, relativamente a qualquer uma das quatro variáveis, um valor inferior a um desvio padrão relativamente a média nacional, o que significa a não conformidade com todas as características enunciadas no conceito de distrito industrial (o único caso que não satisfaz este critério é o concelho da Maia, com o valor de $-1,31$ na variável COEFESP, concelho este cuja diversidade do tecido industrial não permite a sua consideração como distrito industrial). Foram também excluídos os concelhos em que o número de estabelecimentos do sector mais representativo seja inferior a 20 unidades, o que evidencia ausência de "massa crítica" em termos da população de empresas existente (concelhos de Belmonte, Manteigas, Batalha, Castanheira de Pêra, Azambuja e Sines)¹⁰. Na tabela 3 apresentam-se os restantes 16 concelhos, indicando os volumes de emprego total em cada concelho (ano de 1991) e o emprego e o número de estabelecimentos na Indústria Transformadora (média dos anos 1990, 1991 e 1992).

Numa análise sintética desta tabela, verificamos que o emprego dos 16 concelhos considerados corresponde a cerca de 13% do total no Continente, valor inferior ao seu peso no emprego na Indústria Transformadora (I.T.), atingindo este quase um terço do emprego em Portugal Continental, neste sector de actividade. É de salientar que todos estes concelhos possuem uma proporção do seu emprego na Indústria Transformadora superior à média. Destaque para o facto de o rácio médio para estes concelhos (57%) ser mais do dobro da média no Continente (24%). Assim, podemos concluir pela existência de uma taxa de industrialização elevada, correspondendo a este sector, em vários casos, uma percentagem do emprego, na respectiva zona, acima de metade do emprego total.

A este facto, é de adicionar a referência à presença de um número significativo de estabelecimentos industriais, valor nunca inferior a 200, atingindo em alguns concelhos valores acima de um milhar, o que demonstra a existência de uma população de empresas significativa em cada concelho. Por outro lado, em termos da sua aglomeração geográfica, note-se que todos os concelhos exibem uma densidade espacial de trabalhadores na indústria acima da média do agregado de referência (10,7 trabalhadores por KM², face à média de 86,5 por KM² dos 16 concelhos).

Em resumo, na tabela 3, é evidente o perfil industrial de todos os concelhos considerados, caracterizados por uma presença significativa de estabelecimentos industriais, espacialmente aglomerados.

Quanto à especialização intra-industrial de cada concelho, na tabela 4 são apresentados o primeiro e segundo sector dominantes, em termos de volume de emprego e seu peso no emprego da Indústria Transformadora do concelho.

¹⁰ A ausência de um número significativo de estabelecimentos no sector melhor representado não exclui a hipótese do concelho pertencer a uma área de expansão, ou de transição, de um distrito industrial.

Tabela 3: Concelhos Seleccionados

Concelhos	Emprego Total em volume e % (a)	Emprego na I.T. em vol. e % (b)	% de Emprego na I.T. (a)/(b)	Nº de Estab. Industriais (c)	Nº de Trab. da I.T. por KM ² (d)
Águeda	23.340 – 0,54%	14.483 – 1,53%	67,9%	417	43,10
Feira	57.118 – 1,45%	29.024 – 3,07%	50,8%	1.261	137,55
Oliv. de Azeméis	33.260 – 0,84%	17.598 – 1,86%	52,9%	847	115,02
Barcelos	51.467 – 1,30%	22.143 – 2,34%	43,0%	1.115	60,50
Guimarães	79.164 – 2,01%	51.972 – 5,50%	65,7%	1.140	201,44
V.N. Famalicão	57.367 – 1,45%	36.450 – 3,85%	63,5%	981	174,40
Covilhã	21.142 – 0,54%	8.609 – 0,91%	40,7%	214	15,68
Alcobaça	24.091 – 0,61%	10.770 – 1,14%	44,7%	365	26,08
Marinha Grande	14.049 – 0,36%	9.944 – 1,05%	70,8%	241	53,46
Felgueiras	24.512 – 0,62%	17.019 – 1,80%	68,4%	462	146,72
Lousada	20.321 – 0,52%	7.980 – 0,84%	39,3%	282	79,01
Santo Tirso	51.899 – 1,32%	33.250 – 3,52%	64,1%	700	159,86
Alcanena	6.305 – 0,16%	4.510 – 0,48%	71,5%	203	35,51
S. J. da Madeira	9.441 – 0,24%	12.806 – 1,35%	35,6%	336	1.829,43
Paços de Ferreira	21.823 – 0,55%	11.011 – 1,16%	50,5%	886	161,93
Paredes	31.945 – 0,81%	11.950 – 1,26%	37,4%	1.029	76,60
Total Distritos	525.234 – 13,31%	299.519 – 31,67%	57,0%	10.478	86,47
Total Continente	3.945.520 – 100%	945.745 – 100%	24,0%	36.956	10,66
Fontes:	(a) - INE - CENSOS 91 (b) e (c) - MTS - Quadros de Pessoal - média dos anos de 1990, 1991 e 1992.				
Notas:	Chamamos a atenção para o facto de serem utilizadas fontes diferentes no cálculo dos rácios apresentados na colunas (% de Emp. na I.T.) e (d).				

Em termos médios, verificamos que os dois ramos mais importantes, em cada concelho, constituem mais de 70% do seu emprego na Indústria Transformadora. É de realçar que o peso dos dois sectores principais nunca é inferior a 50% do emprego na I.T., de cada concelho, atingindo o valor mais elevado na Covilhã (91%). Em seis casos estamos perante situações claras de mono-especialização, isto é, em ramos complementares ou pertencentes à mesma fileira produtiva (*Textil-Vestuário, ou Produtos Metálicos-Material de Transporte*). Noutros casos, a especialização é em ramos industriais com semelhanças significativas em termos do tipo de produto e organização do processo produtivo (p.ex.^o binómio *Calçado-Vestuário*). É, no entanto, de salientar que em sete concelhos aparentemente serão reduzidas as complementaridades entre os sectores dominantes (como exemplos, os ramos *Calçado-Madeira e Cortiça*, no concelho de St.^a M.^a da Feira, ou *Mobiliário de Madeira-Vestuário*, em Paredes e Paços de Ferreira, que se poderia explicar pela sua vizinhança ao Vale do Ave, ou ainda *Cerâmicas-Calçado* em Alcobaça). Assim, concluímos, que apesar da existência de uma estrutura industrial marcadamente especializada, nem sempre esta especialização é no sentido da predominância de apenas um ramo industrial, existindo em cerca de metade dos concelhos considerados uma especialização dual, ou seja, estamos perante estruturas industriais de maior complexidade.

Note-se que os sectores dominantes nestes concelhos, revelam um quase perfeito ajustamento relativamente ao perfil de especialização dos distritos industriais italianos, identificados por Sforzi (1990), quando este autor refere que: “The majority of [italian] Marshallian industrial districts have a dominant manufacturing specialisation in fashionwear industries - i.e. textile, clothing, footwear, leather goods and tanneries - and wooden furniture; a lower number is dominated by metal goods industries, mechanical and electrical engineering.” (p.84).

A conclusão de que os concelhos referidos anteriormente poderiam ser classificados como distritos industriais ficaria incompleta sem uma referencia adicional ao sistema social local, uma vez que a cada sistema produtivo corresponde um dado contexto de suporte territorial, que o estrutura e que condiciona a sua evolução. Assim, a reprodução social e, nomeadamente, a

reprodução da força de trabalho, para além de um conteúdo económico, deve ser apreciada tendo em conta "...o significado das interferências culturais e simbólicas na modelação das relações materiais." (Reis, 1992, p.229).

Tabela 4: Sectores Dominantes e o seu Peso Relativamente ao Emprego Total na I.T.

Concelhos	Sector Dominante (a)	2º Sector Dominante (b)	Emprego em (a) e % Relativamente ao Total de Emp. na I.T.	Emprego em (b) e % Relativamente ao Total de Emp. na I.T.
Águeda	381 – Prod. Metálicos	384 – Mat. de Transporte	5.108 – 35,3%	2.059 – 14,2%
Feira	324 – Calçado	331 – Madeira e Cortiça	10.165 – 35,0%	9.829 – 33,9%
Oliv. de Azeméis	324 – Calçado	322 – Vestuário	8.476 – 48,2%	2.196 – 12,5%
Barcelos	321 – Têxteis	322 – Vestuário	11.020 – 49,8%	6.888 – 31,1%
Guimarães	321 – Têxteis	322 – Vestuário	29.269 – 56,3%	9.974 – 19,2%
V.N. Famalicão	321 – Têxteis	322 – Vestuário	19.416 – 53,3%	8.113 – 22,3%
Covilhã	321 – Têxteis	322 – Vestuário	5.417 – 62,9%	2.436 – 28,3%
Alcobaça	361 – Cerâmicas – Porcelanas	324 – Calçado	4.025 – 37,4%	1.397 – 13,0%
Marinha Grande	362 – Vidro	381 – Prod. Metálicos	4.006 – 40,3%	1.697 – 17,1%
Felgueiras	324 – Calçado	321 – Têxteis	12.886 – 75,7%	1.305 – 7,7%
Lousada	322 – Vestuário	324 – Calçado	5.099 – 63,9%	1.029 – 12,9%
Santo Tirso	321 – Têxteis	322 – Vestuário	15.942 – 47,9%	9.981 – 30,0%
Alcanena	323 – Curtumes	321 – Têxteis	2.607 – 57,8%	1.092 – 24,2%
S. J. da Madeira	324 – Calçado	382 – Máq. não Eléctricas	6.636 – 51,8%	1.048 – 8,2%
Paços de Ferreira	332 – Mob. de Madeira	322 – Vestuário	5.335 – 48,5%	3.758 – 34,1%
Paredes	332 – Mob. de Madeira	322 – Vestuário	7.839 – 65,6%	2.152 – 18,0%
Total Distritos			153.246 – 51,2%	64.954 – 21,7%

Fonte: MTS - Quadros de Pessoal - média dos anos de 1990, 1991 e 1992.

No caso concreto de um distrito industrial, são determinantes na sua formação a existência de uma homogeneidade e flexibilidade social, que garantam, por um lado, a reprodução do sistema (da mão-de-obra, do factor empresarial,...), e que, por outro lado permitam a sua coesão e a regulação local de conflitos de classe. Para estas duas determinantes contribuem particularmente, no caso português, as questões associadas à organização e ocupação do território, em termos demográficos, e à natureza do papel desempenhado pela pequena agricultura.

No primeiro caso, diversos autores apontam como requisito necessário à emergência de um distrito industrial, uma significativa densidade de recursos humanos de forma a garantir a disponibilidade de mão-de-obra e uma intensidade elevada de comunicações, associada a uma rede urbana assente em pequenas e médias cidades, a qual ofereça um certo número de funções terciárias, evitando as deseconomias de aglomeração presentes nas grandes metrópoles (não só em termos de degradação ambiental, mas, fundamentalmente, em termos do custo de determinados factores como a terra ou o trabalho).

A outra determinante é a estrutura agrária assente na pequena agricultura de cariz familiar (Reis, 1985, 1992; Silva, 1988). Como referimos no primeiro capítulo, a manutenção de alguns elementos que caracterizam a sociedade rural de base familiar, tornam-se relevantes na estruturação e afirmação do modelo de industrialização, nomeadamente em termos do seu papel flexibilizador do mercado de trabalho (recurso ao trabalho ao domicílio e a formas de pluriactividade), na manutenção de uma ética de trabalho (revelada por elevadas taxas de actividade), e na própria oferta empresarial, dada a familiaridade com as funções de gestão e cálculo económico.

Os indicadores presentes na tabela 5, embora limitados pela informação disponível, pretendem precisamente ilustrar de forma sintética os dois factores acima citados.

Tabela 5: Alguns Elementos Caracterizadores do Modelo Social

Concelhos	Taxa de Actividade (a)	% Activos do Sex. Fem. (b)	Dim. Méd. das Exp. Agrícolas (c)	Dim. Méd. das Famílias (d)	Densidade Populacional (d)	Taxa de Urbanização (e)
Águeda	49,4%	41,4%	0,75	3,30	131	22,3%
Feira	49,5%	40,9%	1,51	3,50	562	39,5%
Oliv. de Azeméis	50,8%	41,8%	1,30	3,50	437	38,8%
Barcelos	47,3%	44,0%	2,38	4,00	305	7,8%
Guimarães	52,1%	44,2%	2,75	3,80	611	21,3%
V.N. Famalicão	51,8%	43,5%	3,10	3,50	547	23,9%
Covilhã	42,4%	39,2%	7,23	2,80	98	23,6%
Alcobaça	45,7%	37,9%	2,24	3,00	132	32,7%
Marinha Grande	45,9%	37,6%	1,43	3,00	173	100,0%
Felgueiras	49,1%	41,2%	2,14	3,70	442	13,3%
Lousada	48,9%	40,3%	2,47	3,80	421	0,0%
Santo Tirso	52,9%	42,7%	2,94	3,40	493	38,7%
Alcanena	45,2%	37,9%	3,48	2,90	113	0,0%
S. J. da Madeira	52,8%	43,7%	1,73	3,50	2.636	100,0%
Paços de Ferreira	50,6%	38,3%	1,78	3,90	650	13,7%
Paredes	45,6%	32,9%	1,97	3,80	468	40,2%
Méd. Distritos	49,5%	41,4%	2,24	3,58	318	28,5%
Méd. Continente	44,9%	37,8%	7,04	3,08	106	53,8%
Notas:	Dim. Méd. das Exp. Agrícolas = Superfície Agrícola Útil (SAU) / N° de Explorações em cada concelho.					
Fontes:	Taxa de Urbanização: % da população do concelho que reside em freguesias com mais de 5.000 habitantes.					
	(a), (b), (d), (e), e (f): INE – CENSOS 91.					
	(c) – INE – 1989 – (informação retirada da Base de Dados SALES INDEX – 1998).					

Quanto à forma de ocupação do território, destacamos que a generalidade dos concelhos associa uma densidade populacional acima da média do Continente (a única excepção é a Covilhã, embora com um valor próximo da média), a uma taxa de urbanização reduzida¹¹. Estes indicadores evidenciam, assim, territórios que não sendo marcadamente rurais ou urbanos, associam um grau elevado de industrialização com uma desconcentração da população (muito embora a densidade de comunicações seja intensa), permitindo a coexistência da indústria com a pequena agricultura de base familiar. De facto, a estrutura agrária destes concelhos assenta claramente na pequena exploração, tendo estas uma dimensão média claramente abaixo da média (2,2 ha contra 7 ha).

Esta estrutura agrária, para além de determinar a forma e o modo de ocupação do espaço, é também um dos elementos que condiciona a organização e a estrutura da instituição familiar, ao permitir estratégias familiares de pluriactividade e de diversificação das fontes de rendimento. Estamos perante uma forte mobilização familiar para o trabalho, patente nas elevadas taxas de actividade, com uma significativa participação da mão-de-obra feminina, associada a uma dimensão média das famílias elevada. De facto, neste conjunto, cerca de dois terços dos concelhos, possuem famílias de dimensão média acima dos 3,5 elementos (nomeadamente os concelhos localizados na Região Norte), claramente acima da média no Continente (3,08). Estes dados ajudam à percepção da importância da instituição familiar para o funcionamento da estrutura social local, tal como sugere Reis (1992, p.232).

¹¹ Os casos da Marinha Grande e de São João da Madeira, com taxas de urbanização de 100%, são casos particulares de concelhos constituídos unicamente por duas freguesias (no primeiro caso) ou por uma única freguesia (no segundo caso).

4. Conclusão

A metodologia proposta mostrou-se adequada para a identificação dos distritos industriais marshallianos no caso português. Nos 16 concelhos seleccionados encontramos quer estruturas económicas quer sociais típicas de sistemas produtivos locais, definidos operacionalmente por Sforzi (1990), e descritos por Reis (1992, p. 121), como "...uma entidade sócio-territorial caracterizada pelas interacções internas de um sistema de pequenas e médias empresas envolvidas em diferentes fases de um mesmo processo de produção, espacialmente concentrado, muito ligado à população local e partilhando uma área geográfica restrita".

Referências

- [1] Becattini, G. (1979), "Dal settore industriale al "distretto industriale". Alcune considerazioni sull'unità di indagine dell'economia industriale", *Rivista di Economia e Politica Industriale* 1: Jan. 35-48.
- [2] Becattini, G. (1990), "The Marshallian industrial districts as a socio-economic notion", in F. Pyke, G. Becattini e W. Sengenberger (eds.), *Industrial Districts and Inter-Firm Co-Operation in Italy*. Geneva: International Institute for Labour Studies.
- [3] Beguin, H., 1979, *Méthodes d'Analyse Géographique Quantitative*. Paris: LITEC.
- [4] Bianchi, G. (1998), "Requiem for the Third Italy? Rise and fall of a too successful concept", *Entrepreneurship & Regional Development* 10: 93-116.
- [5] Brandão, A., Pires, A. e Portugal, J. (1998), "Agrupamentos de concelhos de Portugal Continental e sua caracterização", *Revista de Estatística* 1: 71-94.
- [6] Cerejeira da Silva, J.C. (1999), *Distritos Industriais em Portugal Identificação e Avaliação das Externalidades Dinâmicas Associadas*, Dissertação de Mestrado. Braga: Universidade do Minho.
- [7] Cliff, A., Haggett, Ord, J., Bassett e Davies (1975), *Elements of Spatial Structure* Cambridge: University Press.
- [8] Cooper, L. (1963), "Location-allocation problems", *Operations Research* 11: 331-343.
- [10] Everitt, B. (1980). *Cluster Analysis*. New York: Halsted.
- [11] Garofoli, G. (1983), "Le aree sistema", *Politica ed Economia*, XIV: 57-60.
- [12] Garofoli, G. (1994), "Os sistemas de pequenas empresas: Um caso paradigmático de desenvolvimento endógeno", in G. Benko e A. Lipietz (eds.), *As Regiões Ganhadoras*. Oeiras: Celta Editora.
- [13] Gomes, P., Bacelar, S. e Saleiro, E. (1994), "Contributo para a definição de uma tipologia sócio-económica dos concelhos da Região Norte", *Revista de Estatísticas e Estudos Regionais do INE*, 5: 6-15.
- [14] Leandro, A., Esteves, C., e Castro, E. (1997), "O estudo dos sistemas produtivos locais: Reagrupamento dos códigos da CAE segundo o conceito de cluster", *Estudos Regionais*, Direcção Regional do Centro: 31-48.
- [15] Milligan, G. (1980), "An examination of the effect of six types of error perturbation of fifteen clustering algorithms", *Psychometrica* 45: 325-342.
- [16] MacQueen, J. (1967), "Some methods for classification and analysis of multivariate observations", *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability* 1: 281-297.

- [17] Murray, A. (1998), "Assessing clustering methods for exploratory spatial data analysis", comunicação apresentada ao *38th European Congress of the Regional Science Association*, (Vienna: 26-30 Agosto).
- [18] Murray, A., e Estivill-Castro, V. (1998), "Cluster discovery techniques for exploratory spatial data analysis", *International Journal of Geographical Information Science* 12.
- [19] Reis, E. (1997), *Estatística Multivariada Aplicada*. Lisboa: Edições Sílabo.
- [20] Reis, J. (1985), "Modos de industrialização, força de trabalho e pequena agricultura - para uma análise da articulação entre a acumulação e a reprodução", *Revista Crítica de Ciências Sociais* 15: 16-117.
- [21] Reis, J. (1992), *Os Espaços da Indústria: a Regulação Económica e o Desenvolvimento Local em Portugal*. Porto: Edições Afrontamento.
- [22] Sforzi, F. (1987), "L'identificazione spaziale" in G. Becattini (ed.), *Mercato e Forze Locali: Il Distretto Industriale*. Bologna: Il Mulino.
- [23] Sforzi, F. (1990), "The quantitative importance of Marshallian industrial districts in the Italian economy", in F. Pyke, G. Becattini e W. Sengenberger (eds.), *Industrial Districts and Inter Firm Co-Operation in Italy*. Geneva: International Institute for Labour Studies.
- [24] Silva, M.R. (1988), "Developpement industriel et espaces productifs locaux: etude sur la Vallée de l'Ave (Portugal)", *Investigação - Economia* 13, Faculdade de Economia do Porto.
- [25] Silva, M. R. e Figueiredo, O. (1992), "Analyse comparative de deux systèmes productifs locaux dans le Nord Littoral du Portugal", comunicação apresentada ao seminário internacional *Organisation de la Production et Territoire: Les Systèmes de Petits et Moyennes Enterprises*, Pavia, Itália, Julho.
- [26] Sokal, R., e Sneath, P. (1963), *Principles of Numerical Taxonomy*. São Francisco: W.H. Freeman.
- [27] Wise, S., Haining, R. e Ma, J. (1997), "Regionalisation tools for the exploratory spatial analysis of health data", in M. Fischer e A. Getis, (eds.), *Recent Developments in Spatial Analysis - Spatial Statistics, Behavioural Modelling and Computational Intelligence*. Berlin: Springer-Verlag.