

DEMOGRAFIA E ACESSIBILIDADES
-EXPLORAÇÃO DAS RELAÇÕES ENTRE DECLÍNIO POPULACIONAL E
TRANSPORTES NO INTERIOR PORTUGUÊS-

Paulo Rui Anciães

SINOPSE

Neste trabalho procuram-se, através de processos de *data mining*, relações significativas entre o declínio populacional e o isolamento geográfico das freguesias do interior de Portugal no período de 1981-2001, tendo em conta o investimento e desinvestimento na infra-estrutura de transportes ocorrido no mesmo período. Pretende-se isolar o efeito da acessibilidade a centros de atracção populacional através da infra-estrutura de transportes do efeito da acessibilidade à própria infra-estrutura.

A análise sugere que mais do que a acessibilidade à infra-estrutura ou a Lisboa e Porto, a distância a centros locais de atracção populacional tem um papel significativo na explicação dos padrões demográficos apresentados por cada freguesia. Esta conclusão pode no entanto estar ligada a insuficiências dos dados usados e a premissas teóricas feitas no âmbito dos métodos para os tratar.

PALAVRAS-CHAVE

Declínio populacional, isolamento geográfico, infra-estrutura de transportes

1. DESCRIÇÃO E ABORDAGEM AO PROBLEMA

Na maior parte das pequenas localidades no interior de Portugal houve ao longo das duas últimas décadas uma clara tendência de declínio populacional, fenómeno com efeitos negativos na coesão e desenvolvimento sustentado do país (Carrilho *et al* 1993).

Actuando como entrave ao desenvolvimento económico local, o isolamento geográfico é um factor frequentemente apontado como responsável pelo saldo migratório negativo por detrás dos padrões demográficos referidos (Diniz 1998, Pacheco (2001). Estes podem então estar ligados à construção de auto-estradas e vias rápidas e ao encerramento de diversas linhas de caminho de ferro nas regiões do interior durante o período referido (**Fig. 1**).

Figura 1: Alterações na infra-estrutura de transportes no interior português no período 1981-2001

	Estradas fechadas construídas	Linhos de caminho-de-ferro de passageiros encerrados
Nordeste	■ IP4 (Porto-Vila Real-Bragança-Es.)	<ul style="list-style-type: none"> ■ Linha do Douro (Pocinho-Barca de Alva) ■ Linha do Sabor (Pocinho-Duas Igrejas/Miranda) ■ Linha do Tua (Mirandela-Bragança) ■ Linha do Corgo (Vila Real-Chaves)
Beiras	<ul style="list-style-type: none"> ■ IP5 (Aveiro-Viseu-Guarda-Es.) ■ IP3 (Viseu-Coimbra) ■ IC 12 (Nelas-Santa Comba Dão) 	<ul style="list-style-type: none"> ■ Linha do Dão (Santa Comba Dão-Viseu) ■ Linha do Vouga (Espinho-Viseu)
Alentejo	<ul style="list-style-type: none"> ■ A2/A6 (Lisboa-Évora-Es.) ■ A2 (Lisboa-Albufeira/Quarteira) 	<ul style="list-style-type: none"> ■ Linha do Sul-Ramal de Montemor (Torre da Gadanha-Montemor-o-Novo) ■ Linha do Sul-Ramal de Moura (Beja-Moura) ■ Linha de Évora (Évora-Vila Viçosa) ■ Linha de Évora-Ramal de Reguengos (Évora-Reguengos de Monsaraz) ■ Linha de Évora-Ramal de Mora (Évora-Mora) ■ Linha de Évora-Ramal de Portalegre (Estremoz-Portalegre) ■ Linha de Sines (Poceirão a Pinheiro) ■ Linha de Sines-Ramal de Sines (Ermidas/Sado-Sines)

A hipótese do presente estudo é que alterações na infra-estrutura de transporte podem explicar o fenómeno da migração não só através do seu efeito na distância de viagem (em tempo) a centros de atracção populacional (Lee 1969, Dorigo e Tobler 1983) mas também por si só, como factor de repulsão (ou não) ligado a um conceito geral de acessibilidade de cada localidade.

Pretende-se então compreender o papel das alterações ocorridas nas infra-estruturas de transporte entre 1981 e 2001 no comportamento demográfico das regiões afectadas, isolando esse papel da influência da distância, em termos de tempo de viagem, para as áreas de atracção e da localização geográfica das áreas de repulsão.

A abordagem seguida tem várias etapas. Em primeiro lugar identificam-se *outliers* nos dados com base em análise de *clusters*. Após a eliminação destes dados, são procuradas dependências entre as variáveis em estudo e partidos os dados de acordo com as dependências encontradas em cada subconjunto dos dados. Para cada partição são então procuradas relações

lineares entre as variáveis. Foi usado o *software* Polyanalyst 4.6TM para todas as operações.

2. ESCOLHA DAS VARIÁVEIS E PRÉ-PROCESSAMENTO

Define-se como área de estudo uma selecção de NUTS III no interior do país (Fig. 2). A unidade de análise é a freguesia, utilizando-se os dados dos Censos de 1981 e 2001.

Figura 2: Área de estudo (NUTS III)

Norte	Trás-os-Montes e Alto Douro, Douro, Dão-Lafões, Beira Interior Norte, Serra da Estrela, Beira Interior Sul, Pinhal Interior Norte, Pinhal Interior Sul
Sul	Alto Alentejo, Alentejo Litoral, Alentejo Central, Baixo Alentejo

Definem-se como variáveis dependentes a variação percentual da população por freguesia, a diferença entre a variação percentual da população por freguesia e a variação percentual média da população no conjunto de freguesias com a mesma vila mais próxima¹ e ainda a diferença entre a variação percentual da população por freguesia e a variação percentual média da população nas freguesias pertencentes à mesma NUTS III.

Distinguem-se três tipos de variáveis explicativas. As variáveis quantitativas são as distâncias em tempo, obtidas através de modelação em SIG da rede de transportes², de cada localidade a Lisboa ou Porto, à cidade mais próxima e à vila mais próxima, em 2001, e ainda a variação da distância a Lisboa e Porto entre 1981 e 2001. As variáveis binárias são referentes ao facto da localidade estar em 2001 a menos de 10 minutos de uma estação de caminho de ferro, a menos de 20 minutos de uma entrada em auto-estrada ou via rápida e de ter perdido uma estação de caminho de ferro a menos de 10 minutos, devido ao encerramento das linhas entre 1981 e 2001. Finalmente, construíram-se também variáveis binárias para a localização das freguesias no Sul do país³ e em cada uma das NUTS III⁴.

Foram excluídas da análise todas as freguesias respeitantes a vilas e cidades⁵ e foram

¹ Daqui para a frente este conjunto será denominado de “área de serviço” da vila

² Para esta modelação, usaram-se dados geográficos do Instituto Geográfico do Exército para o ano de 2000 (Carta itinerária).

³ Entende-se como Sul o conjunto de freguesias a Sul do rio Tejo

⁴ Excepto para a NUTS III com menor declínio médio de população (Cova da Beira).

⁵ Foram mantidas as freguesias rurais englobadas nas cidades por crescimento destas, uma vez que estas podem

somados os valores da população em localidades compostas por mais de uma freguesia.

Para tratar os valores omissos, decidiu-se atribuir às freguesias criadas após 1981, a mesma variação populacional que as freguesias onde a sua área estava integrada em 1981, sendo este valor corrigido pela redução populacional destas freguesias devido à separação de uma parte da sua área para formar a nova freguesia⁶

3. ANÁLISE EXPLORATÓRIA

3.1. Análise de *clusters*

Com vista a detectar freguesias cujo comportamento demográfico é distinto da generalidade das outras freguesias, foi feita uma busca de *clusters* de dados com base em cada uma das variáveis dependentes e grupos de variáveis independentes⁷ (quantitativas, binárias relativas ao transporte e binárias relativas à localização da freguesia).

Foi usada o algoritmo de procura de *clusters LA* (*Localization of Anomalies*), que procura apenas *clusters* significantes⁸. Em cada grupo de variáveis, após a descoberta de um *cluster*, repetiu-se a busca após retirar as variáveis com base nas quais já existem *clusters*.

Consideraram-se apenas os *clusters* onde um dos atributos é a variável dependente. No caso da variação populacional (Fig. 3.1), um *cluster* encontrado corresponde a variações populacionais positivas e superiores a 11.2 %, associada às menores distâncias à cidade (< 15.5 minutos) e a Lisboa ou Porto (< 95.5 minutos). No caso da variação populacional relativa à média da NUTS III (Fig. 3.2), foi encontrado um cluster de 108 pontos com as menores distâncias à cidade e os maiores valores da variável dependente (positiva e superior a 120.2 %).

apresentar padrões interessantes de variação populacional devido à pequena distância à cidade.

⁶ No caso de freguesias formadas a partir de áreas de 2 ou mais freguesias, admite-se igual proporção de agregação de cada área.

⁷ Foram procurados *clusters* separadamente entre cada variável dependente e cada grupo de variáveis explicativas, uma vez que o algoritmo *LA* só apresentaria *clusters* com maior número de atributos para um consideravelmente maior número de dados (a presente análise tem 1627 registos).

⁸ Este algoritmo funciona colocando os pontos de dados num espaço n -dimensional onde n é o número de atributos envolvidos na exploração. Este espaço é então dividido por hiper-planos, movidos pelo espaço para encontrar áreas que contém números excepcionalmente largos de pontos.

Figura 3.1: Análise de Clusters (Variação populacional e variáveis explicativas quantitativas)

VAR POP CID	(-, 15.5)	[15.5, 28.5)	[28.5, +)
(-, -34.6)			
nºPontos	51	38	40
cluster	--	--	--
[-34.5, 11.2)			
nºPontos	73	72	36
cluster	--	--	--
[11.2, +)			
nºPontos	50	78	40
cluster	2	--	--

VAR POP \L/P	(-, 95.5)	[95.5, 123.5)	[123.5, +)
(-, -34.6)			
nºPontos	51	45	51
cluster	--	--	2
[-34.5, 11.2)			
nºPontos	73	53	31
cluster	--	--	--
[11.2, +)			
nºPontos	50	69	50
cluster	2	--	--

卷之三

4

二

2
(

2

;

-

;

-

;

-

{

;

-

{

;

-

#

I

I

C

C

'C

,

{

-

-

,

(

-

-

,

,

-

-

,

,

-

-

,

,

-

-

,

(

-

-

,

,

-

-

,

,

-

-

Figura 3.2: Análise de Clusters

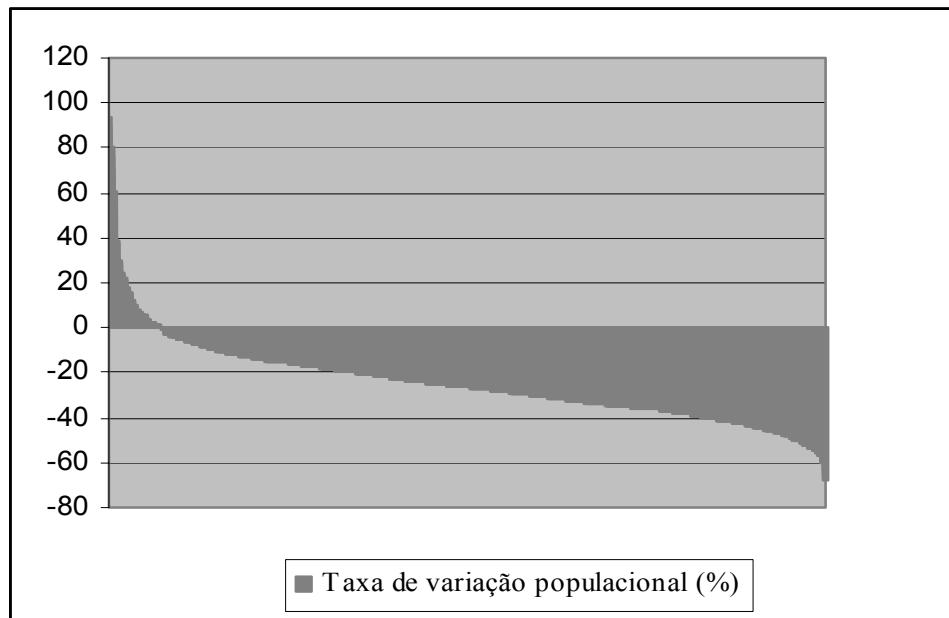
(Diferença da variação populacional em relação à média da NUTS III e variáveis explicativas quantitativas)

DIFNUTS\CID	(-, 15.5)	[15.5, 28.5)	[28.5, +)
(-, -22.2)			
nºpontos cluster	65 --	64 --	53 2
[-22.2, 120.2)			
nºpontos cluster	70 --	56 --	35 --
[120.2, +)			
nºpontos cluster	108 --	68 --	28 --

Estes resultados sugerem que as freguesias com variação populacional positiva ou claramente superior à média da NUTS III a que pertencem⁹ tendem a localizar-se relativamente perto de cidades e de Lisboa ou Porto. Por outro lado, a análise da distribuição dos valores da variação populacional (Fig. 3.3) permite identificar estas freguesias como *outliers* com valores bastante diferentes do resto da população. Como tal, foi decidido excluir da análise as freguesias com variações populacionais positivas no período referido.

⁹ O que representa o mesmo, uma vez que uma diferença de 108 % em relação à média da NUTS III implica uma variação populacional positiva, nos dados existentes.

Figura 3.3. Distribuição dos valores da variação populacional por freguesia



3.2. Procura de dependências

Dependências entre variáveis dependentes e variáveis explicativas

Foi usado de seguida o algoritmo *ARNAVAC*¹⁰ para a identificação das variáveis estatisticamente mais relevantes para a explicação das variáveis dependentes sem precisar a forma específica dessa relação (Fig. 3.4). A vantagem deste método é permitir identificar os subconjuntos dos dados onde essas dependências são válidas.

Figura 3.4: Variáveis mais influentes na previsão das variáveis dependentes

Variável dependente		Variáveis explicativas		
		Quantitativas	Binárias (transporte)	Binárias (localização)
VAR POP	var. mais importantes	VILA, L/P	CF	TMAD
	d.p total	16.18	11.90	11.81
	d.p. subconj.	11.30	8.35	5.85
DIF VILA	var. mais importantes	VILA,CID	CF	PIS, TMAD

¹⁰ Este algoritmo compara a distribuição da variável dependente em hipercubos multidimensionais igualmente povoados.

	d.p total	25.96	28.23	26.78
	d.p. subconj.	16.87	3.99	10.88
DIF NUTS	var. mais importantes	VILA,CID	n-CF	PIS, TMAD
	d.p total	25.75	28.43	26.06
	d.p. subconj.	15.11	11.12	12.57

Notas: d.p total: desvio padrão total da estimativa da variável dependente; d.p.subpop.: desvio padrão total da estimativa da variável dependente no subconjunto de dados onde se verificam as dependências.

Existe uma relação entre variação populacional e a distância à vila mais próxima e a Lisboa ou Porto (**Fig. 3.5**). Esta relação verifica-se para todos pontos a mais de 21 minutos de uma cidade e a mais de 109 minutos de Lisboa ou Porto, devolvendo uma previsão de um declínio de -31.9 % , um valor abaixo da média de -24.8 % para todas as freguesias na área de estudo.

Há também dependências da variação populacional tanto no caso da localização da freguesia na área de serviço de uma estação de caminhos-de-ferro (previsão de -20 %) como no caso da localização fora dessas áreas (devolvendo um valor de – 38.6 %).

A localização fora da NUTS III de Trás-os-Montes e Alto Douro (a segunda região com maior declínio médio de população) é significante e devolve uma previsão de -18.2 %.

Figura 3.5: Dependências entre a variação populacional e as variáveis explicativas

Variáveis quantitativas		Variáveis binárias-transporte			Variáveis binárias-localização			
CIDL/P	(-, 109)	[109, +)	CF	NÃO	SIM	TMAD	NÃO	SIM
(-, 21) previsão	-10.1	-22	previsão	-38.6	-20	previsão	-18.2	-57
nºpontos(célula)	111	53	nºpontos(célula)	1110	517	nºpontos(célula)	1595	32
e.pd(célula)	6.9	14.8	e.pd(célula)	8.9	12.4	e.pd(célula)	9.75	19.5
nºpontos(subpop)	0	0	nºpontos(subpop)	1060	370	nºpontos(subpop)	1566	0
e.pd(subpop)	--	--	e.pd(subpop)	13.8	13.8	e.pd(subpop)	4.9	--
[21, +) previsão	-9.8	-31.9						
nºpontos(célula)	113	105						
e.pd(célula)	8.9	10.9						
nºpontos(subpop)	0	105						
e.pd(subpop)	--	10.9						

No que toca à diferença da variação populacional em relação à média da área de serviço por vila (Fig. 3.6), foram encontradas dependências em relação a ambas as variáveis quantitativas de *input*¹¹. A dependência verifica-se para um valor previsto para a variável dependente sensivelmente na média de toda a área de estudo e é válida apenas para 45 pontos. A variável binária referente à perda de caminho de ferro mantém-se em moldes não muito diferentes do caso da variação populacional absoluta. Verifica-se ainda uma dependência válida para 1271 pontos (77% do total) e com baixo erro padrão para o caso em que a freguesia não se situa nas duas NUTS III com maiores declínios populacionais médios, devolvendo uma previsão de 25 %¹².

Figura 3.6: Dependências entre a diferença da variação populacional em relação à média da área de serviço por vila e as variáveis explicativas

Variáveis quantitativas			Variáveis binárias-transporte			Variáveis binárias-localização		
CID\VILA	(-, 11)	[11, +)	CF	NÃO	SIM	PIS\TMAD	NÃO	SIM
(-, 21)			previsão	-29.4	14.8	NÃO		
previsão	15.2	-34	nºpontos(célula)	1110	517	previsão	25	10.7
nºpontos(célula)	111	70	e.pd(célula)	11.7	9.9	nºpontos(célula)	1496	64
e.pd(célula)	11.3	13.3	nºpontos(subpop)	673	152	e.pd(célula)	10.7	8.3
nºpontos(subpop)	0	45	e.pd(subpop)	11.5	27.79	nºpontos(subpop)	1271	0
e.pd(subpop)	--	10.66				e.pd(subpop)	4.3	--
[21, +)						SIM		
previsão	12.8	-59.1				previsão	-15.1	0
nºpontos(célula)	113	77				nºpontos(célula)	67	0
e.pd(célula)	10.7	8.3				e.pd(célula)	10.2	0
nºpontos(subpop)	0	0				nºpontos(subpop)	0	0
e.pd(subpop)	--	--				e.pd(subpop)	--	--

A diferença da variação populacional em relação à média da NUTS III tem o mesmo tipo de dependências com a variável de perda de caminho de ferro e as duas NUTS III referidas que as outras variáveis dependentes (Fig. 3.7). Há também uma dependência existente para o

¹¹ No caso das variações relativas à área de serviço por vila mais próxima e à NUTS III, não se incluem na análise as variáveis explicativas referentes à distância e à variação da distância a Lisboa ou Porto.

¹² Note-se que o valor médio para a área de estudo da variação populacional relativa ao conjunto de freguesias na mesma área de serviço por vila ou relativa à média da NUTS III é por definição igual a 0.

caso em que as freguesias se situam a menores distâncias para a vila e cidade mais próximas (11 e 21 minutos respectivamente), que devolve um valor de 39 %.

Figura 3.7: Dependências entre a diferença da variação populacional em relação à média da NUTS III e as variáveis explicativas

Variáveis quantitativas		Variáveis binárias-transporte			Variáveis binárias-localização			
CID\VILA	(-, 11)	[11, +)	n-CF	NÃO	SIM	PIS\TMAD	NÃO	SIM
(-, 21) previsão	39	19.1	previsão	17.7	-49.1	NÃO		
nºpontos(célula)	138	74	nºpontos(célula)	1425	202	previsão	35.6	-15
e.pd(célula)	13.2	11.2	e.pd(célula)	10.5	16.6	nºpontos(célula)	1522	67
nºpontos(subpop)	72	0	nºpontos(subpop)	1206	72	e.pd(célula)	10.5	9.7
e.pd(subpop)	92.08	--	e.pd(subpop)	40.84	86.18	nºpontos(subpop)	1336	0
[21, +) previsão	-11.1	-12.2				e.pd(subpop)	50.46	--
nºpontos(célula)	63	42				SIM		
e.pd(célula)	9.6	9.7				previsão	-12.9	0
nºpontos(subpop)	0					nºpontos(célula)	38	0
e.pd(subpop)	--	--				e.pd(célula)	8.7	0
						nºpontos(subpop)	0	0
						e.pd(subpop)	--	--

Partição dos dados com base nas dependências

O algoritmo *ARNAVAC* é também útil para indagar da utilidade de partição dos dados em subconjuntos com vista a analisar posteriormente as diferentes relações existentes entre as variáveis dentro de cada partição, e eventualmente encontrar relações que podem não ser descobertas ao nível do país.

Foi de novo accionado o algoritmo para cada uma das variáveis dependentes e grupos de variáveis explicativas, integrando agora como *input* adicional em todos os casos a variável binária referente à localização da freguesia no Sul do país.

Foram descobertas dependências para o caso da variação relativa à média da área de serviço por vila e existência de caminho de ferro a 10 minutos da freguesia (Fig. 3.8) e para o caso da variação relativa à média da NUTS III e a existência e perda de caminho de ferro (Fig. 3.9). As dependências encontradas existem apenas para o caso do Norte do país e traduzem-se numa previsão de um valor positivo para a variação populacional relativa em casos em que as freguesias têm ou não perderam o serviço de caminho de ferro e num valor negativo para a

variação populacional relativa em casos em que as freguesias não têm esse serviço.

Figura 3.8: Dependências entre a diferença da variação populacional em relação à média da área de serviço por vila, variáveis explicativas e variável binária SUL

SUL\CF	NÃO	SIM
NÃO		
previsão	-22.1	15
nºpontos(célula)	983	416
e.pd(célula)	11.6	9.7
nºpontos(subpop)	485	120
e.pd(subpop)	6.695	24.27
SIM		
previsão	-12.8	-14.1
nºpontos(célula)	127	101
e.pd(célula)	9.5	10.6
nºpontos(subpop)	0	0
e.pd(subpop)	--	--

Figura 3.9: Dependências entre a diferença da variação populacional em relação à média do NUTS III, variáveis explicativas e variável binária SUL

SUL\CF	NÃO	SIM	SUL\n-cf	NÃO	SIM
NÃO			NÃO		
previsão	-29.6	18.2	previsão	9.6	-29.32
nºpontos(célula)	983	257	nºpontos(célula)	257	159
e.pd(célula)	10.7	10.7	e.pd(célula)	10.7	11.9
nºpontos(subpop)	626	97	nºpontos(subpop)	97	42
e.pd(subpop)	10.6	39.45	e.pd(subpop)	39.5	36.42
SIM			SIM		
previsão	-14	-15.9	previsão	5.9	-15.4
nºpontos(célula)	127	58	nºpontos(célula)	58	43
e.pd(célula)	10.1	11.3	e.pd(célula)	11.3	10.1
nºpontos(subpop)	0	0	nºpontos(subpop)	0	0
e.pd(subpop)	--	--	e.pd(subpop)	--	--

Foi testada também a hipótese da população no ano de origem ter alguma influência nas relações encontradas ao nível da freguesia entre as variáveis em estudo. Procuraram-se então dependências entre as variáveis dependentes, grupos de variáveis explicativas e uma variável

referente à população em 1981.

Menores distâncias para vila, cidade e Lisboa ou Porto (respectivamente < 11, <21 e <109 minutos) em localidades com maior população (>612) devolvem uma previsão significativa de -12.3 % no que toca à variação populacional (**Fig. 3.10**). No caso da distância a Lisboa ou Porto, há também uma dependência entre maiores distâncias e maiores populações, que devolve uma previsão sensivelmente igual à média da área de estudo.

Figura 3.10: Dependências entre a variação populacional, variáveis explicativas e população em 1981

POP\VILA	(-, 11)	[11, +)	POP\L/P	(-, 109)	[109, +)
(-, 612)			(-, 612)		
previsão	-10.1	-15.2	previsão	-10.1	-22
nºpontos(célula)	111	70	nºpontos(célula)	111	53
e.pd(célula)	6.9	9.7	e.pd(célula)	6.9	14.8
nºpontos(subpop)	0	0	nºpontos(subpop)	0	0
e.pd(subpop)	--	--	e.pd(subpop)	--	--
[612, +)			[612, +)		
previsão	-12.3	-26.1	previsão	-12.3	-28.9
nºpontos(célula)	214	74	nºpontos(célula)	214	73
e.pd(célula)	28.2	30.8	e.pd(célula)	28.2	25.7
nºpontos(subpop)	42	0	nºpontos(subpop)	42	19
e.pd(subpop)	201.5	--	e.pd(subpop)	201.5	217.4

POP\CID	(-, 21)	[21, +)
(-, 612)		
previsão	-10.1	-9.8
nºpontos(célula)	111	113
e.pd(célula)	6.9	8.8
nºpontos(subpop)	0	0
e.pd(subpop)	--	--
[612, +)		
previsão	-12.3	-32.3
nºpontos(célula)	214	92
e.pd(célula)	28.2	78.7
nºpontos(subpop)	42	0
e.pd(subpop)	201.5	--

No caso da variação populacional relativa à média da área de serviço por vila (**Fig. 3.11**),

as dependências verificam-se também apenas para as maiores freguesias (>612), com distâncias menores a vila ou cidade a implicar previsões de valores positivos e distâncias maiores a implicar previsões positivas.

Figura 3.11: Dependências entre diferença da variação populacional em relação à média da área de serviço por vila, variáveis explicativas e população em 1981.

POP\VILA	(-, 11)	[11, +]	POP\CID	(-, 21)	[21, +]
(-, 612)			(-, 612)		
previsão	15.2	-34	previsão	15.2	-12.8
nºpontos(célula)	111	70	nºpontos(célula)	111	113
e.pd(célula)	11.3	13.3	e.pd(célula)	11.3	10.7
nºpontos(subpop)	0	45	nºpontos(subpop)	0	0
e.pd(subpop)	--	10.66	e.pd(subpop)	--	--
[612, +)			[612, +)		
previsão	34.5	-7.6	previsão	34.5	-19.5
nºpontos(célula)	214	74	nºpontos(célula)	214	92
e.pd(célula)	12.6	8.5	e.pd(célula)	12.6	11.9
nºpontos(subpop)	117	0	nºpontos(subpop)	117	0
e.pd(subpop)	102.1	--	e.pd(subpop)	102.1	--

Para a variação populacional relativa à média da NUTS III (Fig. 3.12), há dependências para populações superiores a 440: distâncias curtas (< 8 minutos para vila e <16 minutos para cidade) implicam previsões positivas, distâncias superiores a 29 minutos para cidade implicam previsões negativas e distâncias intermédias só implicam valor negativo para as maiores populações (> 886).

Figura 3.12: Dependências entre diferença da variação populacional em relação à média da NUTS III, variáveis explicativas e população em 1981.

POP\VILA	(-, 8)	[8, 14)	[14, +)	POP\CID	(-, 16)	[16, 29)	[29, +)
(-, 440)				(-, 440)			
previsão	13.5	14.2	10.5	previsão	13.5	-18.1	-35.2
nºpontos(célula)	33	71	27	nºpontos(célula)	33	51	51
e.pd(célula)	7.5	10.9	8.3	e.pd(célula)	7.5	12.3	6.5
nºpontos(subpop)	0	0	0	nºpontos(subpop)	0	0	0
e.pd(subpop)	--	--	--	e.pd(subpop)	--	--	--
[440, 886)				[440, 886)			
previsão	20.4	-17.6	-10.6	previsão	20.4	10.56	-15.2

nºpontos(célula)	76	88	16	nºpontos(célula)	76	59	154
e.pd(célula)	8.8	11	11.1	e.pd(célula)	8.8	7.5	10.8
nºpontos(subpop)	40	0	0	nºpontos(subpop)	40	0	0
e.pd(subpop)	42.3	--	--	e.pd(subpop)	42.3	--	--
[886,+)				[886, +)			
previsão	38.4	8.3	-9.12	previsão	38.4	-13.7	-19.4
nºpontos(célula)	123	91	17	nºpontos(célula)	123	46	25
e.pd(célula)	12.9	10	7.2	e.pd(célula)	12.8	13.4	10.4
nºpontos(subpop)	81	0	0	nºpontos(subpop)	81	0	0
e.pd(subpop)	106.6	--	--	e.pd(subpop)	106.6	--	--

Com base nos resultados desta secção, criaram-se várias partições dos dados de acordo com a sua localização (Norte ou Sul) ou com população acima ou abaixo do valor mediano da população (**Fig. 3.13**).

Figura 3.13: Partições de dados

	Área	População
P1	Portugal	Qualquer
P2	Portugal	Acima da mediana
P3	Portugal	Abaixo da mediana
P4	Norte	Qualquer
P5	Norte	Acima da mediana
P6	Norte	Abaixo da mediana
P7	Sul	Qualquer
P8	Sul	Acima da mediana
P9	Sul	Acima da mediana

4. PROCURA DE RELAÇÕES

4.1. Relações lineares

Estimaram-se relações lineares para cada variável dependente e partição de dados. Foi utilizado o processo de regressão *stepwise*, escolhendo como modelo final aquele que contém as variáveis mais significativas, a um nível de significância de 20 %.

Variação populacional

Na relações encontradas para esta variável (Fig.4.1), as distâncias à vila e cidade mais próximas são significativas em quase todas as partições e apresentam sempre sinal negativo. No caso do coeficiente para a primeira destas variáveis, em cada zona, é sempre maior para o subconjunto das freguesias com menor população. Os desvios-padrão são geralmente mais baixos para a distância à cidade. A distância a Lisboa ou Porto apresenta coeficientes negativos, mas de menor valor que os da distância para a cidade e vila.

As variações da distância a Lisboa e Porto e a presença de auto-estrada ou via rápida são apenas significantes a 20 % e em muitos poucos casos, o que dá a entender que a construção de auto-estradas no período 1981-2001 não tem grande efeito no comportamento demográfico das freguesias por elas servidas.

A presença de caminho de ferro é no entanto um factor positivo para a variação populacional em cada freguesia. As freguesias com acesso a uma estação de caminhos-de-ferro apresentam quebras populacionais entre cerca de 2 e 3 % menores que as restantes freguesias na mesma partição. Apenas no caso das maiores freguesias do Sul este coeficiente é negativo, o que se pode relativizar dado o grande desvio padrão que esta estimativa apresenta. Nesse modelo, a perda de caminho-de-ferro é significante e tem sinal positivo.

As variáveis binárias de localização são quase todas significantes no Norte mas não no Sul do país. Também não há nenhum efeito aditivo da localização das freguesias no Norte ou no Sul do país. É de notar ainda que o valor da constante do modelo é sempre superior no caso das freguesias de maior população, em cada zona.

As relações encontradas explicam apenas uma pequena parte da variação da variável dependente. De facto, o valor do coeficiente de determinação R^2 é apenas maior que 50 % no caso das freguesias de maior população no Norte do país. As regressões para o país ajustam-se pior aos dados (maior R^2) que as regressões para o Norte mas melhor que as regressões para o Sul. Por outro lado, as regressões nas partições de maior população ajustam-se sempre melhor, qualquer que seja a zona.

Figura 4.1: Relações lineares para a explicação da taxa de variação populacional por freguesia

Variável	P1	P2	P3	P4	P5	P6	P7	P8	P9
R²	0.3	0.46	0.34	0.38	0.56	0.33	0.22	0.27	0.25
e.pd	7,73	9,72	11,36	8,43	9,61	11,46	10,82	8,79	12,03
c	-16.46	-13.99	-19.16	-15.99	-7.4	.19.96	-11.96	-17.47	-7.87**

	(1.28)	(1.9)	(2.04)	(0.85)	(2.17)	(2.15)	(3.43)	(2)	(5.74)
VILA		-0.22 (0.06)	-0.3 (0.07)	-0.1 (0.05)	-0.25 (0.07)	-0.26 (0.08)	-0.24 (0.12)		-0.38* (0.2)
CID	-0.1 (0.02)	-0.19 (0.03)	-0.11 (0.04)	-0.13 (0.02)	-0.19 (0.04)	-0.11 (0.04)		-0.15 (0.05)	
L/P			-0.05 (0.02)		-0.03* (0.02)	-0.05 (0.02)	-0.09 (0.03)		-0.13 (0.05)
VAR L/P	0.03** (0.02)	0.04** (0.03)			0.06** (0.04)				
AE			-1.44** (1.03)						
CF	2.13 (0.56)	2.02 (0.8)	3.15 (1.09)	2.95 (0.63)	3.23 (0.87)	2.7 (1.17)		-4.65 (2.1)	
n-CF								5.05 (2.57)	
SUL									
TMAD	-4.95 (0.85)	-11.13 (1.19)	-4.61 (0.94)	-9.1 (0.76)	-13.91 (1.08)	-4.79 (1.01)	-	-	-
DOU	-3.51 (0.9)	-5.32 (1.32)		-5.13 (0.77)	-9.34 (1.22)		-	-	-
DAO	1.56* (0.84)	3.72 (1.24)	6.15 (1.63)			5.53 (1.78)	-	-	-
BIN	-3.85 (0.94)	-3.73 (1.74)		-6.74 (0.86)	-6.63 (1.59)		-	-	-
SE			8.29 (2.16)	-2.78 (1.24)	-4.78 (1.9)	8.77 (2.33)	-	-	-
BIS	-3.98 (1.45)	-4.2 (1.73)		-5.37 (1.44)	-4.94 (2)		-	-	-
PIN	1.53** (1.09)	2.45 (1.47)					-	-	-
PIS	-3.74* (2.05)	-7.14 (2.12)		-6.01 (1.88)	-8.17 (2.08)		-	-	-
AA		-2.43** (1.61)		-	-	-			
AL			-	-	-	-			
AC			-	-	-	-			
BA	-2.15* (1.28)		-	-	-	-			

Notas

- 1- Quadros com valor dos coeficientes e desvio padrão da estimação (em parêntesis)
- 2- Coeficientes significativos a 5 % excepto onde: * (significativo a 10%) ** (significativo a 20%)
- 3- R²: Coeficiente de determinação; e-pd: Erro padrão da estimação. c: Termo constante

Diferença da variação populacional em relação à média da área de serviço por vila

As relações procuradas para esta variável (Fig.4.2) encontram as mesmas variáveis explicativas que para o caso da variação populacional, exceptuando a distância e a variação da distância a Lisboa ou Porto¹³.

A distância à vila tem coeficiente sempre significativo e sempre negativo. O valor do coeficiente é consideravelmente superior ao caso do modelo anterior. A distância à cidade mais próxima, por seu lado, tem agora uma influência muito reduzida.

A presença da auto-estrada a 10 minutos faz aumentar sempre o declínio populacional, embora essa relação apenas num caso seja significante a menos de 5 %. Por outro lado, a presença de caminho de ferro diminui o declínio excepto nas freguesias no Sul e no subconjunto das maiores freguesias nessa área. Em ambos os casos, a perda de caminho de ferro nas últimas década tem influência positiva na variação populacional.

A variável binária SUL é significativa a 20 % e tem um valor positivo, o que retrata o facto que as freguesias dentro da mesma área de influência de uma vila, são mais homogéneas demograficamente no Sul do que no Norte do país.

Para todas as partições, as relações para esta variável dependente ajustam-se pior aos dados que os modelos para a variação populacional.

Figura 4.2: Relações lineares para a explicação da diferença da variação populacional em relação à média da área de serviço por vila

Variável	P1	P2	P3	P4	P5	P6	P7	P8	P9
R ²	0.33	0.22	0.23	0.28	0.23	0.22	0.22	0.35	0.17
e.pd	8,33	9,94	11,25	9,01	10,03	11,50	9,97	8,76	10,21
c	1.22*	1.87	-0.97	-4.69	1.62	0.14*	2.16**	8.59	-0.66*
	(0.77)	(0.92)	(1.3)	(1.56)	(1)	(1.13)	(2.01)	(2.46)	(2.41)
VILA	-0.22	-0.38	-0.42	-0.23	-0.46	-0.35	-0.31	-0.48	-0.25**
	(0.49)	(0.06)	(0.07)	(0.05)	(0.07)	(0.08)	(0.12)	(0.14)	(0.18)

¹³ Para esta variável dependente a também para o caso da variação populacional relativa à média da NUTS III decidiu-se não se incluir na análise estas duas variáveis explicativas.

CID	0.08 (0.02)	0.05* (0.03)	0.06* (0.03)	0.04** (0.03)	0.05** (0.04)				
AE	-1.14* (0.62)		-1.53** (1.01)	-2 (0.69)		-1.96* (1.05)			
			2.81 (1.08)	1.57 (0.7)		2.73 (1.21)	-3.8 (1.66)	-8.48 (2.08)	
		2.06* (1.06)			2** (1.3)		4.24 (2.1)	6.33 (2.58)	
SUL	1.1** (0.83)			-	-	-	-	-	-
TMAD	-8.25 (0.82)		-3.03 (0.92)	-8.27 (0.83)	-1.9** (1.13)		-	-	-
DOU				-4.1 (0.92)		-	-	-	-
DAO	2.97 (0.83)	2.71 (1.06)	2.7* (1.56)		1.54** (1.15)	-2.79 (1)	-	-	-
BIN	-3.22 (0.91)	-3.28* (1.66)		-4.81 (0.94)			-	-	-
SE	-5.12 (1.53)	-3.44 (1.68)		-7.09 (1.57)			-	-	-
BIS			-6.83 (2.25)	2.73 (1.23)		-5.9 (2.54)	-	-	-
PIN				-7.54 (2.05)	-3.05* (2.12)		-	-	-
PIS	-5.63 (2.16)						-	-	-
AA				-	-	-			
AL	3.54* (1.98)	4.5 (2.06)	16.93 (8.04)	-	-	-	-	-	-
AC				-	-	-	-	-	-
BA				-	-	-	2.29** (1.53)		

Diferença da variação populacional em relação à média da NUTS III

As relações procuradas para esta variável (Fig.4.3) apresentam as mesmas variáveis explicativas que o caso anterior. Também neste caso a estimativa se ajusta sempre pior aos dados que nas relações encontradas para a variação populacional.

A distância à vila mais próxima tem sempre um coeficiente negativo e da mesma ordem

de grandeza que para as outras variáveis dependentes. Neste caso, no entanto, a probabilidade de significância é maior, em cada área, para as mesmas partições dos dados segundo a população. A distância à cidade mais próxima é significante em quase todos os casos, tendo sinal negativo e um valor semelhante ao caso do modelo da variação populacional absoluta.

A variável relativa à existência de saída para auto-estrada ou via rápida não tem neste modelo qualquer papel significativo. A existência de caminho de ferro tem um efeito positivo exceptuando no mesmo subconjunto que é excepção nos modelos anteriores. A perda de caminho de ferro tem também um valor positivo nas maiores freguesias quer no Norte quer no Sul do país, embora apenas significante a 10 %.

O coeficiente da variável binária referente à localização no Sul é positivo no subconjunto das maiores freguesias. Uma freguesia no Sul do país apresenta em média uma diferença de mais 4.12 % em relação à média da NUTS III do que uma freguesia no Norte.

Figura 4.3: Relações lineares para a explicação da diferença da variação populacional em relação à média do NUTS III

Variável	P1	P2	P3	P4	P5	P6	P7	P8	P9
R²	0.43	0.32	0.28	0.37	0.34	0.26	0.23	0.34	0.20
e.pd	10.85	11.5	10.83	9.58	10.84	10.04	5.26	4.51	3.04
c	-3.27 (0.75)	3.96 (1.67)	6.2 (2.04)	-1.45** (1.33)	9.24 (1.89)	9.43 (2.54)	3.23* (1.98)	8.12 (2.7)	1.65* (2.82)
VILA		-0.21 (0.07)	-0.34 (0.07)	-0.1* (0.05)	-0.26 (0.07)	-0.29 (0.08)	-0.32* (0.13)	-0.33 (0.16)	-0.5* (0.21)
CID		-0.09 (0.02)	-0.15 (0.03)	-0.07 (0.03)	-0.12 (0.02)	-0.17 (0.04)	-0.07* (0.04)	-0.1 (0.05)	-0.13 (0.05)
AE									
CF	2.07 (0.56)	1.78 (0.8)	2.98 (1.05)	2.85 (0.64)	2.44 (1.03)			-5.92 (2.11)	
n-CF					2.08** (1.46)			5.04* (2.6)	
SUL		4.12 (1.22)		-	-	-	-	-	-
TMAD	-13.73 (0.81)	-5.92 (1.19)	-4.23 (1.1)	-11.2 (0.82)	-4.17 (1.16)		-	-	-
DOU	-7.23 (0.77)	-2.8 (1.25)		-6.42 (0.86)		3.78 (1.38)	-	-	-
DAO			6.03		2.44	10.12	-	-	-

BIN	-9.51 (0.9)	-9.18 (1.79)	(1.66) (1.38)	-3.47 (0.99)	-8.68 (0.99)	-7.18 (1.62)	(1.19) (1.85)			
SE	3.29 (1.19)				3.03 (1.28)			-	-	-
BIS	-10.4 (1.43)	-11.37 (2.04)	-6.49* (3.54)	-11.85 (1.71)	-10.87 (2.03)			-	-	-
PIN	-7.24 (1.08)	-6.65 (1.5)	3.1** (2.39)	-7.17 (1.16)	-5.67 (1.53)	6.26 (2.54)		-	-	-
PIS	-16.37 (2.04)	-11.58 (2.19)	-7.43 (3.54)	-15.97 (1.93)	-11.83 (2.14)			-	-	-
AA	7.02 (1.25)				-					
AL		5.37 (2.19)	12.63** (8.15)		-					-0.02
AC	-5.55 (1.17)				-					
BA	-9.36 (1.28)	-4.96 (1.69)			-			-4.22 (1.76)	-3.49* (1.93)	

5. ALGUMAS LIMITAÇÕES

Alguns pressupostos teóricos adoptados nesta análise podem ser limitativos. De facto, ao equacionar-se saldo migratório com variação populacional assume-se que o saldo demográfico natural é igual em todas as freguesias. Assume-se também que tanto variáveis fluxo para o período 1981-2001 como variáveis *stock* para 2001 influem nas variações demográficas (uma variável fluxo). Por outro lado, as migrações podem responder com algum hiato temporal aos estímulos ocorridos (Mohlo 1986), pelo que o efeito da infra-estrutura de transporte na demografia pode não ser visível no período em que as alterações ocorrem. Não foram também considerados e diferenciados os factores de atracção de vilas e cidades nem as alterações na rede de estradas nacionais e regionais durante o período em causa.

As dependências e relações encontradas padecem de pobre adequação aos dados, o que não é uma desvantagem se considerarmos que os dados não são uma população mas sim uma amostra¹⁴, sendo então mais importante atender à correcta estimativa dos parâmetros da

¹⁴ Neste caso a população seria o conjunto de variações demográficas em cada freguesia ao longo do tempo

população (Granger e Newbold 1976, Achen 1982). Por outro lado, esta pode ser limitada pela multicolinearidade entre dados como as distâncias à vila e cidade mais próximas.

6. CONCLUSÕES

O isolamento geográfico e o transporte explicam apenas uma parte dos fenómenos demográficos existentes no interior português.

As variáveis mais significativas na explicação do declínio populacional absoluto e relativo são a distância à vila e à cidade mais próxima, especialmente em freguesias com menor população no início do período em estudo. A construção de auto-estradas e vias rápidas e a respectiva redução na distância de cada localidade para as duas grandes cidades portuguesas parece não estar relacionada com a variação populacional nas áreas servidas. A presença de estações de caminho de ferro na actualidade é positiva no Norte mas negativa no Sul do país, tendo a perda das estações tem o comportamento inverso a este.

De um modo geral, o investimento na infra-estrutura de transporte rodoviária não parece ser um factor que consiga travar a tendência de declínio populacional nas regiões do interior.

REFERÊNCIAS

- Achen, C H (1982) *Interpreting and Using Regression*. Sage Publications. Beverly Hills, California.
- Carrilho, M J. et al (1993) Alterações demográficas nas regiões portuguesas entre 1981-1991., in *MPAT/MA - O Programa de Desenvolvimento Regional e o Mundo Rural*. MPAT/MA, Lisboa.
- Diniz, F S. (1998) *Desenvolvimento Rural, Que Perspectivas?: O caso de Alto Trás-os-Montes e Douro*. Fundação D. Afonso Henriques, Lisboa.
- Dorigo, G., Tobler, W R. (1983) Push-pull migration laws. *Annals of the Association of American Geographers*. **73**, 1-18.
- Granger, C e Newbold, P. R² and the transformation of regression variables. *Journal of Econometrics*. **4**, 205-10.
- Lee, T R. (1969) A theory of migration. *Demography* **3**, 47-57.
- Mohlo, I I. (1986) Theories of migration: a review. *Scottish Journal of Political Economy*. **33(4)**, 396-419.
- Pacheco, E. (2001) Alteração das Acessibilidades e Dinâmicas Territoriais na Região Norte: Expectativas, Intervenções e Resultantes. Tese de Doutoramento (policopiada). Faculdade de Letras da Universidade do Porto.