

# Análise Espacial de BD Paroquiais: antes e depois da Fusão

Fernanda Faria\*  
NEPS

Pedro Rangel Henriques  
DI

Universidade do Minho  
20 de Fevereiro de 2004

## Resumo

O objectivo deste artigo é mostrar as potencialidades da exploração dos dados demográficos recolhidos nas paróquias segundo o método de Norberta Amorim, MRP, através de um sistema de Informação Geográfica que permita analisar o comportamento social numa paróquia ou entre paróquias de uma zona ao longo dos séculos.

Tal exploração pressupõe a aquisição dos dados que se extraem dos registos de baptizado, casamento e óbito e o seu arquivo em bases de dados normalizadas, trabalhando individualmente com as BD paroquiais e uma BD central que segue a mesma estrutura.

## 1 Introdução

A Demografia Histórica (DH) é um dos campos de investigação da História e estuda os principais fenómenos demográficos num espaço temporal de algumas centenas de anos até à actualidade. Em Portugal, esse estudo pode ser feito a partir de meados do século XVI, dado que a sua principal fonte de informação encontra-se nos registos paroquiais (RP) — Registos de Baptizado, Registos de Casamento e Registos de Óbito — cujo início da sua sistematização ocorreu por essa altura.

Os RP serviram de base à Metodologia de Reconstituição de Famílias, desenvolvida por Louis Henry e Michel Fleury, e posteriormente à Metodologia de Reconstituição de Paróquias (MRP) concebida por Norberta Amorim, professora e investigadora do Núcleo de Estudos da População e Sociedade (NEPS). Esta última metodologia permite a reconstituição de uma paróquia pela identificação de cada indivíduo e das famílias com que se relaciona através do cruzamento dos dados constantes nos três tipos de RP [Amo91]. A partir da reconstituição da vida dos indivíduos e das famílias é possível analisar indicadores demográficos de Natalidade, Fecundidade, Nupcialidade, Mortalidade e Mobilidade.

É na sequência da cooperação entre o NEPS e docentes e investigadores do Departamento de Informática (DI) da mesma universidade que surge a idealização do Sistema para o Estudo da Evolução Demográfica (SEED), que pretende dar

---

\*Bolseira da FCT ao abrigo do Programa SAPIENS 99

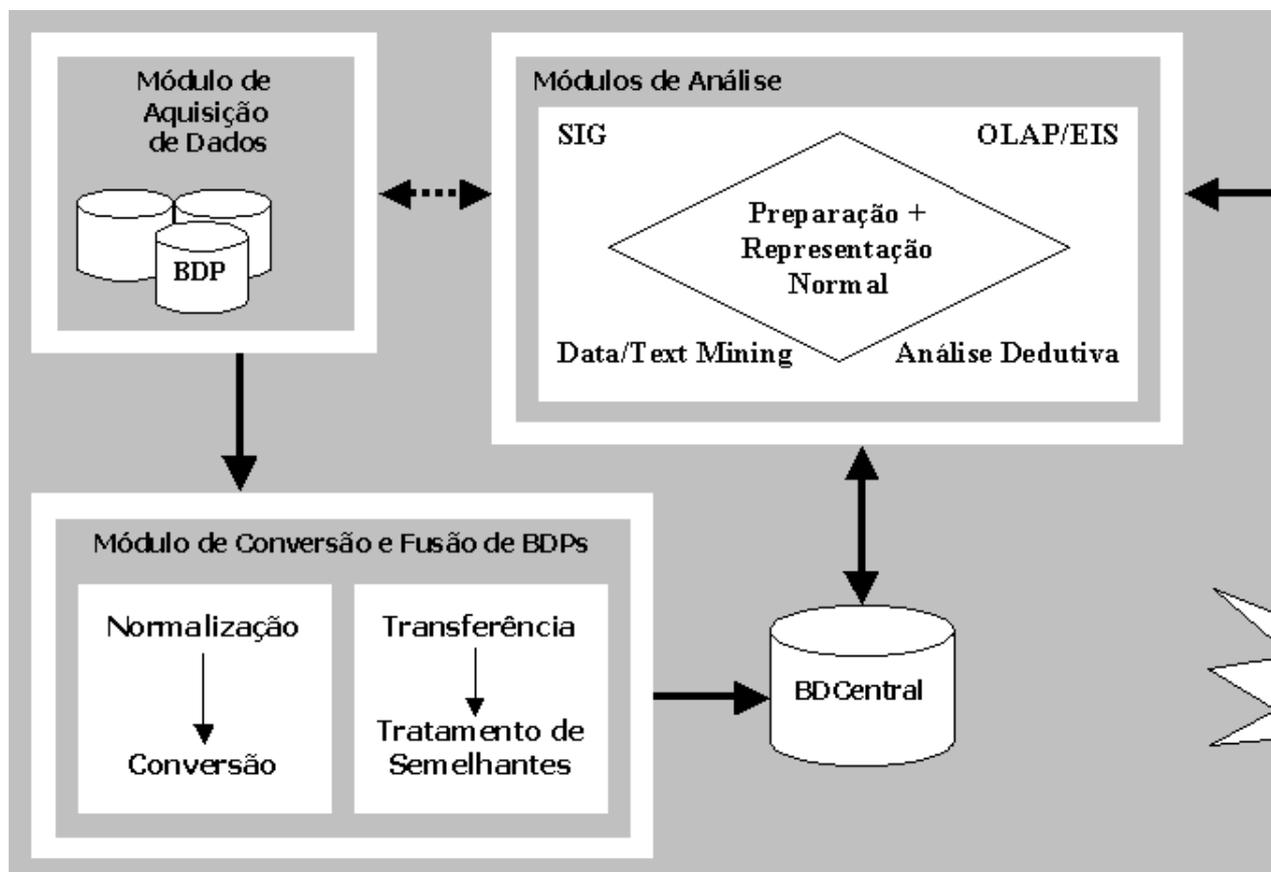


Figura 1: Arquitectura do Sistema SEED

suporte informático aos processos de Análise Demográfica dos dados recolhidos segundo o MRP. Na figura 1 é apresentada a arquitectura do SEED que compreende todos os processos desde a aquisição até à divulgação dos dados [Rod00].

No esquema da figura 1 identificam-se 5 camadas:

- Camada de Aquisição, inclui a interface para introdução e validação das fichas de indivíduos e famílias (após o levantamento dos RPs) e o sistema de armazenamento (BDPs).
- Camada de Fusão, destinado a fazer a adição dos dados de uma paróquia à BDC identificando indivíduos semelhantes e resolvendo colisões;
- Camada de Armazenamento Normalizado, consiste no próprio repositório de dados normalizados de forma a poderem ser utilizados pela camada de Análise, concentrando as diversas paróquias estudadas;
- Camada de Análise, para a realização de diversos tipos de análise sobre a informação obtida dos RPs eventualmente cruzada com outras fontes

documentais;

- Camada de Gestão do Conhecimento, para criação e manutenção de um repositório do conhecimento adquirido, incluindo as funções de armazenamento, verificação-validação e divulgação.

A exploração dos dados, questão central desta comunicação, é realizada recorrendo a ferramentas EIS de Análise Multidimensional (OLAP) como o TM1 ou o SAS, ferramentas de Análise Estatística como o SPSS, ou ferramentas de *Data Mining* como o Clementine; geralmente os resultados são visualizados em forma de tabela ou gráfico. Mas pela natureza dos RP esta informação contém uma referência de localização pelo que muitas análises podem ser visualizadas de uma outra perspectiva — Perspectiva Georeferenciada — geralmente fornecida por Sistemas de Informação Geográfica (SIG). Os resultados das análises obtidas dos dados georeferenciados são visualizados num mapa através de cores permitindo uma visão mais nítida e directa de todo o conjunto em estudo.

Este artigo tem como principal finalidade a reflexão sobre as potencialidades dos SIGs aplicados à DH e ao MRP em concreto, a uma paróquia ou a um conjunto de paróquias.

A próxima secção faz uma breve descrição dos SIG e apresenta as possibilidades da sua aplicação a um sistema de dados de indivíduos e famílias; enumeram-se estudos que se podem fazer ao nível da paróquia e faz-se referência a estudos interparoquiais. Entretanto, e ainda nessa secção, referem-se o tipo de fontes e suas relações que tornam possíveis essas análises geo-espaciais.

Na terceira secção, apresentam-se as aplicações do SEED que já estão desenvolvidas e em teste — o SRP que engloba todo o processo de aquisição dos dados para carregamento das Bases de Dados Paroquiais (BDP); o ConFusBDP que é responsável pela conversão (normalização) e fusão das BDPs na Base de Dados Central (BDC) — que vão permitir o estudo mono e multiparoquial georeferenciado. Serão também apresentados alguns resultados obtidos de simulações realizadas com um SIG específico (Geomedia), onde se fizeram, entre outros, análises de mobilidade.

A última secção apresenta as conclusões e algum trabalho futuro que interesse desenvolver.

## 2 Exploração Demográfica Georeferenciada

Por georeferenciação, entende-se a atribuição de uma componente localizacional à nossa informação. Estima-se que 80% da informação manipulada no nosso dia-a-dia possui uma referência espacial. Como informação geograficamente referenciada ou georeferenciada entende-se como informação relativa a qualquer entidade susceptível de ser localizada no espaço através de coordenadas geográficas. As entidades podem ser visíveis ou palpáveis, como por exemplo, edifícios, ruas e árvores. Podem ser também não palpáveis, como sejam limites territoriais (distritos, freguesias,...) e acontecimentos (acidentes, crimes, incêndios, ...). Ou seja, informação georeferenciada é aquela que é relativa a algo que possua localização espacial.

A maior parte da informação tem implícita uma componente espacial que permite localizá-la no espaço: um endereço, um país ou uma cidade e um SIG permite a sua manipulação, visualização e análise de uma forma que não é possível

com a informação em formato tabular ou estatístico. O armazenamento de informação num SIG difere do armazenamento de informação numa qualquer base de dados no facto de a informação possuir uma referência espacial. Enquanto uma base de dados armazena informação que simplesmente contenha referência a uma dada localização, isto é, possui um endereço, o SIG utiliza a georeferenciação como principal forma de armazenar e aceder à informação.

Os principais conceitos de um SIG são:

- **Visualização** — fornece uma nova perspectiva de visualização das entidades espaciais e dos resultados das análises efectuadas;
- **Criação de Mapas** — permite criar mapas personalizados possibilitando ao utilizador indicar: as entidades espaciais a ver (por exemplo, colocar visível as paróquias e as igrejas e colocar invisível os concelhos, as ruas e os imóveis); a simbologia a adoptar (por exemplo, as igrejas representadas por um ponto laranja, as paróquias representadas com fundo amarelo e limite azul marinho); e a escala a usar na produção do mapa;
- **Consultas topológicas** — reconhece e analisa relações topológicas entre os vários elementos representados num mapa através da determinação de condições:
  - de *adjacência* (Que freguesias se encontram adjacentes à freguesia S. José de S. Lázaro?);
  - de *conteúdo* (Que ruas se encontram na freguesia de S. Vítor?);
  - de *proximidade* (Que freguesias se encontram num raio de 10 km do rio Este?);
  - e de *intersecção* (Que ruas cruzam com a Av. da Liberdade?)
- **Mapas temáticos** — permite efectuar todo o tipo de mapas temáticos que consistem em apresentar os elementos espaciais coloridos com uma cor determinada pelos valores que têm em relação a determinado critério ou em determinado atributo. Considerando um exemplo, descreve-se de seguida como se cria um mapa temático.

Uma Câmara Municipal pretende implementar um sistema de incentivos à natalidade nas suas freguesias. Para isso, pretende estudar a natalidade em cada uma delas a partir do número de nascimentos.

Inicialmente identificamos quais os intervalos de valores a usar na nossa análise; supondo que se utilizariam 3 intervalos, estes poderiam ser:

- até 10 nascimentos;
- de 11 a 25 nascimentos;
- a partir de 26 nascimentos.

De seguida, indica-se a cor correspondente a cada intervalo definido, podendo ser um *dégradé* de tons de uma só cor (várias tonalidades de uma cor, começando pela tonalidade mais clara até a tonalidade mais acentuada) ou um conjunto de várias cores. No nosso exemplo, consideremos para cada um dos intervalos de valores as seguintes cores:

- 1ª intervalo: Vermelho;

- 2º intervalo: Laranja;
- 3º intervalo: Verde.

O mapa temático é então criado colorindo as freguesias do concelho tendo em conta o número de nascimentos e o intervalo de valores em que se encontram. As freguesias com menos de 11 nascimentos estão pintadas a vermelho, as freguesias que tiveram entre 11 e 25 nascimentos encontram-se pintadas a laranja e as restantes a verde.

Outros exemplos de mapas temáticos podem ser: mapa da natalidade em Portugal no ano de 2002, mapa dos partidos eleitorais eleitos por concelho, mapa dos restaurantes de Lisboa por tipo de cozinha (Tradicional, Francesa, Italiana, Chinesa, Espanhola, ...);

- **Padrões de Comportamento** — potencia o aparecimento de padrões de comportamento dada a perspectiva visual que fornece [San01]. Tomando o exemplo dos incentivos à natalidade referido no ponto anterior, a visualização do número de nascimentos por freguesias numa perspectiva geográfica permite, por exemplo, verificar que as freguesias situadas a oeste do concelho têm uma taxa de natalidade mais baixa ou que as freguesias em que o rio passa têm uma natalidade mais alta, e a partir dessas observações procurar as suas causas. Adicionando outras fontes de dados ao mapa, podem ser visualizadas características que possam ser a causa ou origem das observações atrás referidas. Por exemplo, se se visualizar o rendimento *per capita* médio das famílias por freguesia, poderia ser constatado que as famílias residentes nas freguesias situadas a oeste do concelho têm efectivamente um menor rendimento que as famílias residentes na outra ponta do concelho o que pode constituir uma causa para a baixa taxa de natalidade nessa zona.

Também nos RP, encontramos informação que é possível representar no mapa. No tratamento dos registos de baptismo, existe informação sobre o local de baptismo e podemos encontrar informação sobre a naturalidade e residências dos pais e padrinhos. Ao tratarmos os registos de casamento, é possível localizar o casamento e, por vezes, a naturalidade e a residência dos nubentes, das suas famílias e dos padrinhos. No caso dos registos de óbito, além do local do enterro, podemos ter referida a naturalidade e a última residência do falecido.

Os elementos georeferenciados obtidos no levantamento dos RP correspondem a paróquias, localidades, concelhos, países, etc. Estes locais podem ser hierarquizados da seguinte forma: um país contém vários distritos, um distrito contém vários concelhos, um concelho contém várias paróquias, uma paróquia contém vários lugares, e um lugar contém vários sítios ou ruas — possibilitando o estudo georeferenciado a qualquer nível na hierarquia de locais.

Como é sabido, na DH não são os RPs que nos interessam isoladamente, mas sim a informação sobre o Indivíduo e sobre a Família que a partir deles se reconstrói, para então se calcularem variáveis que são indicadores do comportamento geral da sociedade em diferentes meios e épocas. Assim, a partir dos RPs passa a ser possível georeferenciar o trajecto de vida do indivíduo e das famílias a que este está associado e, portanto, abre-se a possibilidade de associar os indicadores sociais calculados a posições geográficas.

Importa agora identificar o valor acrescentado que esse facto — *a capacidade de*

*localizar os valores médios de fecundidade, nupcialidade, mortalidade, etc. sobre um mapa* — nos pode aportar, na medida em que as facilidades resultantes de associar a BDParoquial ou Central a um SIG se traduza essencialmente numa forma diferente de ver a informação.

Para representar num mapa uma determinada variável, por exemplo a fecundidade, o que se faz é escolher uma cor numa escala em que se associam cores a intervalos de valores diferentes. Portanto se no período de tempo em estudo se pintar a paróquia P com a cor C e ao lado se apresentar a legenda com o dito código de cores, o ganho que se obtém é pequeno: uma leitura rápida do valor calculado (obtido com igual facilidade pela simples leitura do valor absoluto ou percentagem).

Este efeito visual é particularmente propício ao estabelecimento de comparações. Assim, se se colocarem lado a lado mapas com a paróquia pintada de cores diferentes em diferentes períodos de tempo ou em relação a indicadores diferentes, já o ganho é maior pois este estudo comparativo torna-se mais simples/intuitivo (mas talvez não tão intuitivo como em um comum gráfico de barras).

É óbvio que as vantagens desta forma de análise crescem se se estudarem em simultâneo várias paróquias — nesse caso é possível mostrar sobre o mesmo mapa as diferentes paróquias pintadas a cores diferentes permitindo assim, numa rápida inspeção, a extracção de resultados comparativos entre o comportamento de determinadas zonas relativamente a um parâmetro em estudo. Torna-se pois evidente que os benefícios de um SIG aplicado à BD do SEED são distintos consoante se trabalhe numa situação mono ou multiparoquial.

Por exemplo, o estudo da taxa de natalidade de uma paróquia, analisado através de um SIG, corresponde a ver a paróquia em estudo com determinada cor. Essa cor isolada não nos diz nada. Mas se se comparar várias paróquias, cada uma com a sua cor dependendo do intervalo de valores em que a natalidade de cada uma se encontra, já se obtém uma percepção do significado das cores de cada paróquia, isoladamente ou em conjunto. Se, por exemplo, pudéssemos criar um mapa temático das várias paróquias de um concelho, no que toca à taxa de fecundidade ilegítima, poderíamos facilmente detectar a que tinha uma taxa menor — representada por uma cor de tonalidade mais suave — e a que tinha a taxa mais elevada — representada por uma cor de tonalidade mais forte.

Para a maioria das variáveis tradicionais, o estudo monoparoquial só faz sentido se for para comparar resultados em diferentes épocas. Nomeadamente, o estudo da Mobilidade só pode ser feito sobre a BD Central onde existam várias paróquias.

Para além desta análise imediatista de representação geográfica dos indicadores demográficos tradicionais, a DH abre outras hipóteses quando se relaciona informação existente nos RP com outras fontes, como registos prediais, róis de confessados, ... Seguem-se alguns exemplos de análises que nesse caso podem ser efectuadas por um SIG:

- Genéricas:
  - Longevidade versus Valor Patrimonial ou Dimensão da propriedade;
  - Idade ao casamento versus Valor Patrimonial;
  - Distribuição das casas pelo número de elementos do agregado familiar;
  - Distribuição das localidades da paróquia por taxas de mortalidade

- Distribuição dos terrenos pelo tipo de cultura;
- ...
- Específicas:
  - Distribuição das casas com rendimentos superiores a 30 réis anuais;
  - Acompanhamento dos membros de determinada família e seus descendentes relativamente às sucessões e aquisições;
  - Distribuição das propriedades do Sr. António na paróquia;
  - Localização dos terrenos com mais de 4000 metros quadrados e nome dos respectivos proprietários;
  - Distribuição por herdeiros, não necessariamente descendentes, dos bens imóveis legados pelo falecido Sr. José;
  - Identificação dos indivíduos residentes numa área de 250 metros da igreja paroquial;
  - Distribuição dos bens imóveis dos habitantes elegíveis e respectivo valor patrimonial;
  - Localização dos terrenos de cultivo da paróquia;
  - ...

Sendo possível a ligação das fontes documentais cruzadas às entidades geográficas, podem-se efectuar todo o tipo de análises espaciais não possíveis com ferramentas de análise comuns. Por exemplo, a listagem das propriedades do Sr. António não fornece a mesma perspectiva que a visualização da sua distribuição. Esta última perspectiva fornece à primeira vista a ideia da quantidade, da proximidade, da adjacência e das áreas das propriedades, bem como o enquadramento espacial destas no conjunto de propriedades da paróquia.

## 2.1 Exploração Multiparoquial Geo-referenciada

Um SIG quando aplicado a análises multiparoquiais permite efectuar todas as que são possíveis e aquelas que se revelaram sem interesse num âmbito monoparoquial.

Assim, a visualização de análises de natalidade, nupcialidade, fecundidade e mortalidade fornecem uma visão geral de determinado conjunto de paróquias ou outro limite administrativo de nível superior facilitando claramente todos os estudos comparativos.

É possível efectuar comparações entre paróquias do mesmo concelho ou concelhos diferentes, é possível efectuar a comparação entre uma paróquia e o concelho a que pertence, etc.

No entanto, deve-se realçar a possibilidade do estudo de um fenómeno fundamental da DH impossível de visualizar num contexto monoparoquial — a **Mobilidade** — pois, pela sua natureza, envolve sempre no mínimo um local de origem e um local de destino. A análise espacial da Mobilidade não é, de forma nenhuma, tão linear como a análise dos outros fenómenos geralmente estudados pela DH. Foram estudadas técnicas de representação deste fenómeno por um SIG. São elas:

- **Por cores:** O local de origem encontra-se colorido com uma cor (por exemplo, verde) e os locais de destino encontram-se coloridos com várias tonalidades de outra cor (por exemplo, vermelho). As várias tonalidades do vermelho indicariam o fluxo migratório entre os dois locais — origem e destino — correspondendo a tonalidade mais clara aos menores fluxos e a tonalidade mais escura aos maiores fluxos de migração.
- **Por setas:** A base da seta indica a origem do indivíduo a outra extremidade indica o destino de migração. O fluxo migratório pode ser visualizado como uma etiqueta (*label*) junto à seta correspondente ou na própria grossura da seta — isto é, uma seta mais fina indica um menor fluxo e uma seta mais grossa indica um maior fluxo de migração.

### 3 Aquisição, Armazenamento e Fusão de RPs

Após se ter discutido as diferentes formas em que a ligação de um SIG a um sistema de dados de indivíduos e famílias (baseado, à partida na informação recolhida dos RP, mas prevendo já o seu cruzamento com outras fontes que nos informam sobre bens imóveis e outras propriedades e pertences do indivíduo) pode, em geral, ajudar o processo de análise, pretende-se nesta secção concretizar a proposta tendo por suporte o sistema de informação construído segundo o MRP com base no SEED. Para tanto e antes de ilustrar as ideias com os resultados de estudos georeferenciados (como a mobilidade) vai-se referir o mecanismo de aquisição e construção das BDPs (SRP) e o módulo de conversão e fusão com resolução dos conflitos resultantes da ocorrência de indivíduos semelhantes (ConFusBDP).

#### 3.1 O SRP

O SRP supõe um processo prévio manual de levantamento dos dados dos RP e de reconstrução do indivíduo e família, fornecendo então toda a funcionalidade necessária à criação de uma BDP, isto é, suporta as tarefas de inserção, manutenção e cruzamento dos dados de um indivíduo ou de uma família.

Esta aplicação foi desenvolvida tendo sempre em vista os seguintes aspectos:

1. A aplicação deve manter o *workflow* usado pelos historiadores demógrafos (HD) no processo manual ou nas aplicações desenvolvidas anteriormente, como por exemplo a criação de uma família implica a associação de pelo menos um indivíduo a essa família.
2. A aplicação deve funcionar tendo como principal objectivo a correcção dos dados em detrimento da rapidez da sua introdução na BDP.

Relativamente ao primeiro ponto, os HDs podem utilizar a aplicação usando as suas funcionalidades de acordo com a sua forma habitual de trabalhar. Assim: podem variar a ordem de levantamento dos RPs — começar pelos registos de baptismo ou pelos de casamento; podem variar a forma de inserção dos dados — por exemplo, (1) inserir todos os indivíduos e criar depois as famílias com os indivíduos já existentes ou (2) criar uma família e ao mesmo tempo criar os indivíduos a ela associados.

Quanto ao segundo ponto, a aplicação obriga a que qualquer inserção, alteração ou eliminação seja explicitamente confirmada antes de ter efectivamente lugar na BDP de forma a evitar incongruências nos dados. Obviamente, a necessidade de confirmar / cancelar todas as operações traduz-se em tempo acrescido na inserção de cada RP, mas garante a correcção dos dados.

Foram construídos dois formulários principais que tratam da inserção e manutenção das principais entidades do MRP — Indivíduo e Família — e diversos outros formulários para tratamento de informações relacionadas com as duas entidades anteriormente referidas (informações sobre as residências, profissões, herdeiros, sufrágios, ...) ou para tratamento de informação auxiliar (listas de locais, listas de profissões, graus de parentesco, tipos de missas, ...). Far-se-á de seguida uma breve descrição dos principais formulários do SRP, visto que manipulam os dados fundamentais e básicos de qualquer RP que constituem a base das análises. Por ser o principal objecto de estudo da DH no contexto paroquial começa-se pela apresentação do formulário de Família ilustrado na figura 2. Este formulário

**Família**

N.º Família: 8      Situação Família: Legítima      Família: 2 de 5515

Tipo Família: 2      Famílias das quais conhecemos registo de baptismo de todos os filhos, data de casamento, data de óbito do primeiro cônjuge falecido e desconhecemos a data de nascimento da mulher.      Início Obs.: 13-06-1748      Fim Obs.: 29-07-1756

Pai: 16959      Alberto Lino Melo

Mãe: 19368      Maria Ana Ressurreição

C A S A M E N T O

Data: 13-02-1741      Data Fixa?       Local: Lajes do Pico -F- 460102      interVeniEntes

Pai: Idade:      Estado civil:      N.º Casamentos:

Mãe: Idade:      Estado civil:      N.º Casamentos:

4 filhos

Nind	Nome	D. Nascimento	D. Obito
21	José	11-11-1741	
22	Francisco	02-12-1744	
23	Luís Inácio Melo	13-06-1748	
24	Indeterminado	25-03-1751	

Residências      Obs.:

Localizar      Inserir      Alterar      Eliminar      Fechar

Figura 2: SRP – Formulário de Família

manipula os dados específicos da entidade Família indicando o(s) indivíduo(s) que a constituíram, a sua classificação segundo o tipo e situação de família, as datas de início e fim de observação, e observações que não se enquadrem nos campos disponíveis. Caso exista o registo de casamento desta família, os dados sobre a data e local de casamento, a idade e número de casamentos anteriores de cada nubente são também registados. É também possível obter uma listagem dos filhos originados pela família em questão apresentando os seus nomes, datas de nascimento, de casamento, de fim de observação e de óbito. Sempre que a

operação pretendida seja a associação ou consulta dos indivíduos de uma família, a aplicação redirecciona-nos para o formulário de Indivíduo ilustrado na figura 3. O formulário de Indivíduo apresenta a informação específica do indivíduo, como

The screenshot shows a software window titled 'Indivíduo'. At the top, there are fields for 'Nº Ind.' (19368) and 'Nome' (Maria Ana Ressurreição). To the right, there are fields for 'Indivíduo' (18543) and 'de' (20466). Below these are 'Tipo Ind.' (4) and 'Sexo' (M). There are also fields for 'Família de Origem', 'Título', and 'Alcunha'. A section titled 'nascimento' contains 'Data Nasc.', 'Data Fixa?', 'Local Nasc.' (Lajes do Pico - F- 460102), and 'Filiação'. Below this is a section for 'Informação dos pais ao nascimento' with fields for 'Pai' and 'Mãe' (Idade and Estado civil). There are also 'Início Obs.' and 'Fim Obs.' fields. On the right side, there are buttons for 'Profissões', 'Residências', and 'Assinaturas'. At the bottom, there are navigation buttons (back, forward) and action buttons (atribuir, localizar, inserir, alterar, eliminar, fechar).

Figura 3: SRP – Formulário de Indivíduo

datas de início e fim de observação, a sua classificação e observações. Optou-se pela visualização dos seus dados organizados por três separadores, cada um relativo a cada tipo de RP — Baptismo, Casamento e Óbito — de forma a dar uma perspectiva mais organizada sobre os seus actos vitais.

O primeiro separador apresenta os dados relativos ao nascimento, como pode ser visto na figura 3, que são data e local de nascimento, filiação, idade e estado civil dos seus pais à época do nascimento.

O segundo separador é ilustrado na figura 4 e apresenta a listagem das relações familiares em que esse indivíduo se pode encontrar, isto é, as famílias legítimas ou ilegítimas que esse indivíduo constituiu.

Na figura 5 encontra-se ilustrado o terceiro separador que apresenta a informação sobre o óbito do indivíduo. Podem ser consultados ou editados dados como data e local de óbito, causa de morte, classificação paroquial, entre outros.

Todos os outros formulários da aplicação podem ser acedidos a partir dos dois principais atrás referidos.

### 3.2 O ConFusBDP

Para se proceder a estudos multiparoquiais, devem ser considerados os seguintes processos:

N.º Família	Situação	Data de Casamento
8	Legítima	13-02-1741

Figura 4: SRP – Separador das Relações Familiares do Indivíduo

Figura 5: SRP – Separador da informação sobre o Óbito

1. **Conversão** Já existem diversas BDPs com formato distinto das BDPs obtidas através do SRP. Para que exista coerência e uniformização nos dados usados nas análises de DH, estas têm de ser convertidas para o formato de BDP usado actualmente — SRP.
2. **Fusão** Dado que o levantamento dos RPs é feito paróquia a paróquia, existem tantas BDPs quantas as paróquias tratadas. As BDPs devem ser aglutinadas numa só, pois o correcto tratamento analítico impõe que não haja duplicação de indivíduos e de famílias. As situações de duplicação dos indivíduos e famílias, quando não devidamente tratadas, ocorrem inevitavelmente dada a possibilidade de mobilidade entre paróquias.

De forma a auxiliar o utilizador nos dois processos referidos, foi desenvolvida uma aplicação — o ConFusBDP — que tomando uma BDP efectua a sua conversão, quando necessário, e a sua fusão para a BDC.

O processo de Conversão de BDPs é feito de forma automática depois de ter identificado manualmente a codificação a usar nos campos relativos a locais e profissões para estar de acordo com o formato usado no SRP.

O processo de Fusão de BDPs é efectuada entre BDPs com o formato SRP. Este processo pretende colocar numa só BD — na BDC — os indivíduos e famílias de várias paróquias, uma paróquia de cada vez, o que leva à possibilidade de se encontrarem indivíduos e famílias que já existem na BDC. Este processo efectua-se em quatro passos, são eles:

1. Transferência dos Indivíduos
2. Tratamento dos Indivíduos Semelhantes
3. Transferência das Famílias
4. Tratamento das Famílias Semelhantes

No primeiro passo, a aplicação procura na BDC indivíduos ou famílias com características muito semelhantes aos que se querem inserir; no caso de não encontrar semelhantes faz a sua inserção directa; no caso de encontrar semelhantes regista esse facto e coloca em *stand-by* a sua inserção.

Terminado o passo de inserção directa dos indivíduos, procede-se ao tratamento dos Indivíduos Semelhantes. O utilizador percorre todos os indivíduos semelhantes de um indivíduo a inserir para aferir se algum é efectivamente o mesmo. Se decidir que o indivíduo que pretende transferir corresponde a um indivíduo semelhante, faz a actualização dos dados do último com os dados do primeiro. Se não encontrar um indivíduo semelhante que o possa considerar o mesmo que o indivíduo a inserir, procede então à sua inserção na BDC. Na figura 6 é apresentado o formulário de Tratamento de Indivíduos responsável por apoiar o utilizador neste passo.

**Indivíduo encontrado** Tipo Ind. 3 Início Obs. Fim Obs.

3 Isabel

**Datas**

Nind	Nascimento	Casamento	Obito
3	19-07-1733		12-02-1785

**Locais**

Nind	Nascimento	Casamento	Obito
3	460102		460102

**Foram encontrados 2 Indivíduos Semelhantes**

Nind	Nome	Tipo Ind.	Início Obs.	Fim Obs.
619	isabel st.andre	0	07-05-1745	09-10-1769
710	isabel francisca	0	14-09-1780	14-09-1780

**Datas**

Nind	Nascimento	Casamento	Obito
619			

**Locais**

Nind	Nascimento	Casamento	Obito
619			

Buttons: << < > >> Inserir Actualizar Fechar

Figura 6: ConFusBDP — Tratamento de Indivíduos Semelhantes

Tanto no primeiro passo como no segundo, sempre que é inserido um indivíduo na BDC, são também transferidos os seus dados relativos a profissões, residências e assinaturas.

O terceiro passo é efectuado da mesma maneira que o primeiro mas relativamente às famílias; ou seja, procura na BDC famílias semelhantes à família que se pretende inserir. Posteriormente no quarto passo é feito o tratamento das famílias semelhantes de forma idêntica ao tratamento dos indivíduos semelhantes.

Visto que em alguns dos passos do processo da Fusão, nomeadamente no segundo e quarto passo, é necessária a intervenção do utilizador este processo não é automático mas semi-automático.

### 3.3 Alguns Resultados obtidos com um SIG

Utilizando o SIG Geomedia da Intergraph, foram já feitas algumas simulações com BDPs de algumas paróquias da ilha do Pico, Açores.

Um primeiro exemplo de resultado obtido com um SIG é a criação de um mapa temático por determinado critério. Neste caso, tomando a taxa de natalidade, as paróquias são pintadas na cor correspondente ao intervalo de valores em que se encontram. Como se pode ver na figura 7, as paróquias usadas na análise encontram-se pintadas a determinada cor; essa cor indica qual o intervalo de valores em que as paróquias se enquadram. A legenda apresenta a escala de cores e valores usada na representação visual do resultado da análise, começando numa cor mais clara para o intervalo de valores mais baixos até a cor mais escura para o intervalo de valores mais elevados.

A percepção do resultado analítico obtido quando se utiliza um SIG é mais rico que a apresentação desses resultados em forma tabular (folha de cálculo), pois fornece uma visão conjugada do valor da taxa de natalidade com as áreas das paróquias, com as suas situações geográficas na ilha e com os factores proximidade e adjacência entre elas.

Como já foi referido na secção 2, um outro exemplo do que se pode obter com um SIG é um estudo de mobilidade que está ilustrado na figura 8. O que se pode ver é um fluxo migratório no ano de 1720, representado pela linha, entre duas paróquias — Origem: Sta Luzia; Destino: Criação Velha — e um valor associado a esse fluxo — neste caso, 3 pessoas saíram de Sta Luzia para Criação Velha.

## 4 Conclusão

Neste artigo pretendemos mostrar algumas das potencialidades de um SIG aplicado à DH, referindo alguns exemplos em que este traz mais-valias às análises fundamentais desta área de estudo e outros exemplos em que não existem vantagens na sua utilização. Em conclusão, este tipo de análise deve ser feito quando os dados manipulados por um SIG podem fornecer uma perspectiva espacial intuitiva e elucidativa, não fornecida por outra ferramenta de apresentação de resultados, caso dos estudos comparativos de variáveis espaciais (áreas, posições, etc) e da mobilidade.

Foram também apresentadas as aplicações desenvolvidas para o suporte à aquisição e armazenamento electrónico dos RPs — o SRP, Sistema de Reconstituição de Paróquias, e o ConFusBDP, Sistema de Conversão e Fusão de BDPs — constituindo as camadas iniciais da arquitectura SEED e a base para qualquer tipo de exploração e divulgação dos dados.

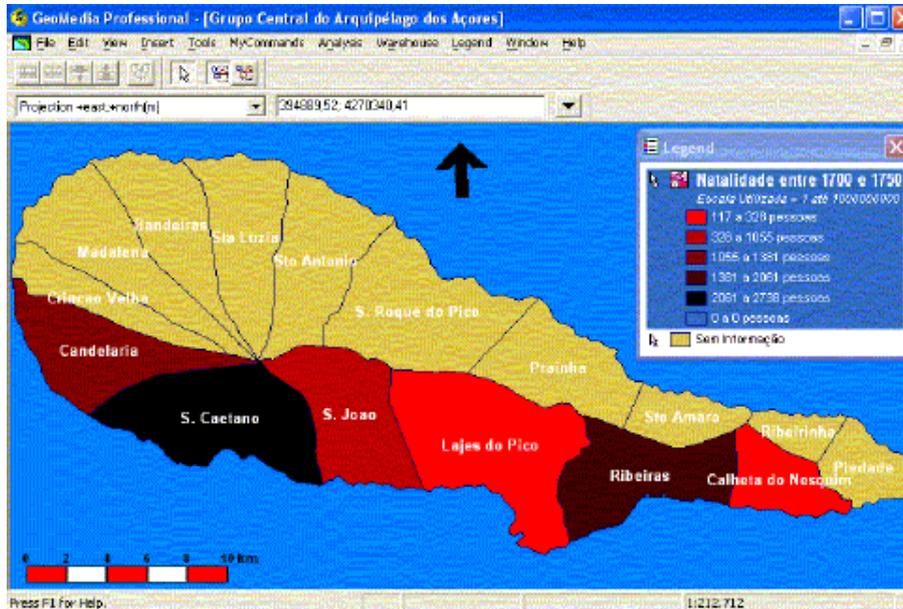


Figura 7: Mapa temático sobre a natalidade em algumas paróquias do Pico

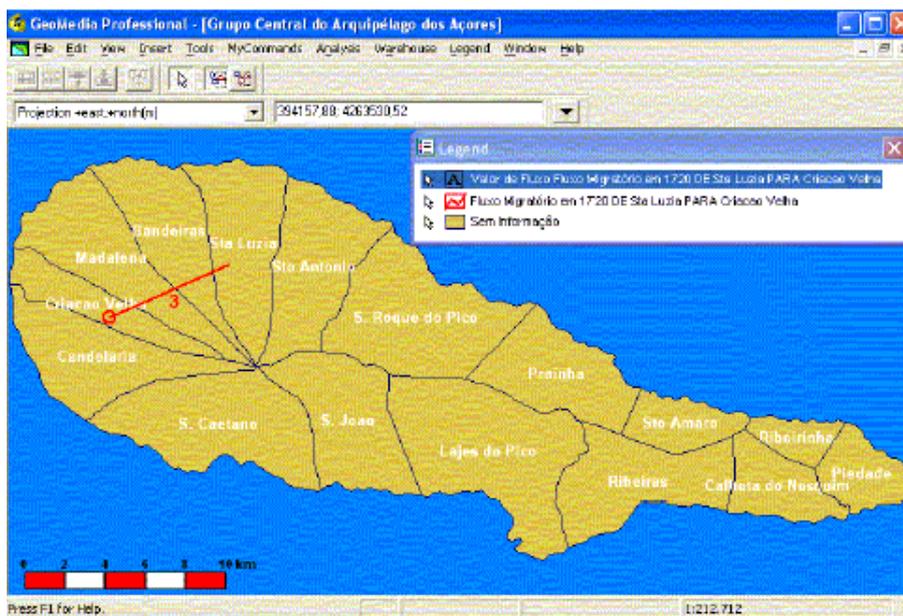


Figura 8: Exemplo de Representação da Mobilidade

Uma das linhas de trabalho futuro está relacionado com o suporte à aquisição; pretende-se que o sistema permita introduzir (talvez usando técnicas de anotação estrutural de documentos) os registos na íntegra e ajude (de forma semi-automática) a fazer o seu cruzamento com vista à reconstituição paroquial.

## Referências

- [Amo91] Maria Norberta Amorim. *Evolução Demográfica de Três Paróquias do Sul do Pico*. Instituto de Ciências Sociais, Universidade do Minho, 1991.
- [Rod00] Fátima Rodrigues. *Arquitectura Heterogénea para Extracção de Conhecimento a partir de Dados*. PhD thesis, Universidade do Minho, Dezembro 2000.
- [San01] Maribel Santos. *PADRÃO — Um sistema de Descoberta de Conhecimento em Bases de Dados Georeferenciadas*. PhD thesis, Universidade do Minho, 2001.