



Boletim de Ciências Geodésicas

ISSN: 1413-4853

bcg\_editor@ufpr.br

Universidade Federal do Paraná  
Brasil

PHILIPPI CAMBOIM, SILVANA; ROBBI SLUTER, CLÁUDIA  
ESTUDO SOBRE UM ALGORITMO PARA A CONSTRUÇÃO DE MAPAS DE PONTOS  
DE CONTAGEM

Boletim de Ciências Geodésicas, vol. 19, núm. 1, enero-marzo, 2013, pp. 65-83  
Universidade Federal do Paraná  
Curitiba, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=393937728005>

- Como citar este artigo
- Número completo
- Mais artigos
- Home da revista no Redalyc

redalyc.org

Sistema de Informação Científica  
Rede de Revistas Científicas da América Latina, Caribe, Espanha e Portugal  
Projeto acadêmico sem fins lucrativos desenvolvido no âmbito da iniciativa Acesso Aberto

# ESTUDO SOBRE UM ALGORITMO PARA A CONSTRUÇÃO DE MAPAS DE PONTOS DE CONTAGEM

*A Study About an Algorithm to Build Dot Maps*

SILVANA PHILIPPI CAMBOIM  
CLÁUDIA ROBBI SLUTER

Universidade Federal do Paraná  
Departamento de Geomática – SCT  
Centro Politécnico – Jardim das Américas  
81.531-990 Curitiba Paraná - Brasil  
silvanacamboim@ufpr.br; robbi@ufpr.br

## RESUMO

Os mapas de pontos de contagem são importantes ferramentas para análise da densidade espacial de fenômenos. Na sua construção, é estipulada uma quantidade N de ocorrências para que sejam representados por um ponto. O objetivo deste estudo foi propor um método para agrupar os dados por pontos, um algoritmo para automatizar a definição destes agrupamentos e a verificação da potencialidade do método e do algoritmo produzindo um exemplo com os dados de casos de AIDS (Síndrome da imunodeficiência adquirida) no Brasil. Os dados utilizados foram provenientes do Programa Nacional de DST (Doenças Sexualmente Transmitidas) e AIDS, que abrangem os casos por município de 1980 a 2010. Foi gerada uma série multitemporal de mapas, usando para orientar o agrupamento a divisão hierárquica do espaço territorial brasileiro utilizada pelo IBGE (Instituto Brasileiro de Geografia e Estatística): Município – Mesorregião – Microrregião – Unidade da Federação. Também são detalhadas as principais fases na elaboração deste tipo de mapa, apresentando as melhores práticas quando já estabelecidas na literatura em cartografia e também algumas tendências atuais e lacunas na pesquisa sobre este tópico. O algoritmo se mostrou uma ferramenta eficiente para automatizar o processo de geração do mapa com pontos distribuídos de forma geograficamente embasada, proporcionando a visualização da distribuição do fenômeno ao longo do tempo.

**Palavras-chave:** Mapas de Pontos de Contagem; AIDS; Análise Multitemporal.

### ABSTRACT

Dot maps are important tools to analyze the spatial density of a phenomenon. During the process of map design, a certain quantity  $N$  of occurrences is established to be represented by a point. The objective of this study was to create a method to group data by point, an algorithm to a grouping definition and the assessment of both method and the algorithm producing an example using the cases of AIDS (Acquired Immunodeficiency Syndrome) in Brazil, with data of the National Program of STD (Sexually Transmitted Diseases) and AIDS, that enclose cases from 1980 to 2010, to generate a series of multi-temporal maps, using as a guide in the process of grouping points the hierarchic division of the Brazilian territorial space used by IBGE (Brazilian Institute of Geography and Statistics). Also the main steps on designing this type of map project are detailed, presenting best practices, when already established in cartography literature, and also some current trends and gaps in research on this topic. The algorithm proved to be an efficient tool to automate the process of generating dot maps distributed in a geography-based way, and propitiate the visualization of the phenomenon through time.

**Keywords:** Dot Maps; AIDS; Multi-temporal Analysis.

### 1. INTRODUÇÃO:

Mapas de pontos de contagem são, segundo definição dada por Slocum (2009), aqueles nos quais se define uma certa quantidade ( $N$ ) de algum fenômeno para ser representada por um ponto e, a seguir, localizam-se os pontos nos locais onde o mesmo tem a maior probabilidade de acontecer.

Neste estudo, o fenômeno escolhido foi o número de casos de AIDS no Brasil por município, compilados e disponibilizados pelo Programa Nacional de DST e AIDS, do Ministério da Saúde (MINISTÉRIO DA SAÚDE, 2012), para os anos de 1980 a 2010. Os dados são provenientes do Sistema Nacional de Agravos de Notificação (SINAN), Sistema de Controle de Informações Laboratoriais (SISCEL) e Sistema de Informações sobre Mortalidade (SIM).

O propósito do mapa é fornecer uma ferramenta para análise da distribuição dos 608.230 casos de AIDS notificados no país de 1980 a 2010, em um ambiente interativo e utilizando a animação como forma de apresentar a evolução dos casos através dos anos. Epidemiologistas tradicionalmente usam mapas para examinar a distribuição espacial de doenças (por exemplo: Eng, 1999). Os mapas de pontos de contagem facilitam a identificação de focos de casos, a formulação de hipóteses sobre as causas de infecção, os fatores de risco espacialmente distribuídos e a análise dos dados. Internacionalmente, os casos de AIDS têm sido particularmente mapeados desta maneira (CDC, 2008; DSHS, 2008) em decorrência da rápida dispersão da doença a partir dos raros casos registrados antes do início da década de 1980, ainda que não utilizasse otimizações para a construção dos mapas.

Durante o projeto cartográfico de um mapa por pontos de contagem destacam-se as seguintes questões a serem consideradas:

- Determinação da região em que cada ponto será posicionado,
- Definição do tamanho do ponto e do valor representado por ele e
- Alocação dos pontos dentro das regiões selecionadas.

Estas decisões estão intimamente relacionadas com aquelas a serem tomadas nas etapas iniciais do projeto cartográfico, como a escala do mapa e o propósito do mesmo. Neste tipo de mapa, o importante não é como são posicionados individualmente os pontos, mas sim, o padrão de densidade formado pelos mesmos.

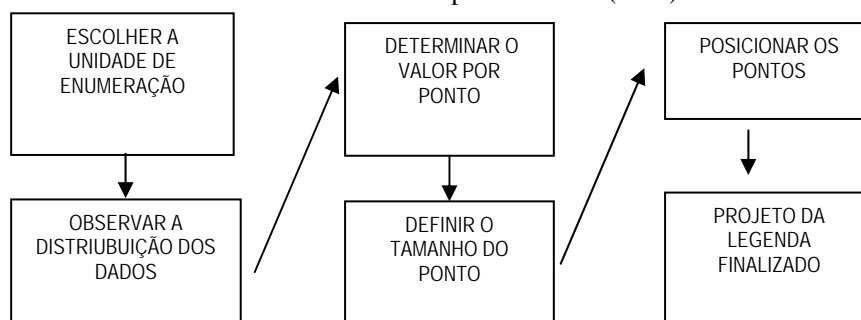
## 2. METODOLOGIA

### 2.1 Unidades de Enumeração e Distribuição de Dados

Para execução de um mapa de pontos de contagem deve ser observada a sequência apresentada no diagrama abaixo (Figura 1).

Figura 1 - Esquema com os passos para a criação de um mapa de pontos de contagem.

Fonte Adaptada: JIANG (2004).



No início do projeto do mapa é necessária a escolha da unidade de enumeração, que define o grau de detalhamento para o mapa final. Por exemplo, neste estudo, foram escolhidos como Unidade de Enumeração os estados da federação. A escolha foi realizada com base na escala do mapa a ser produzido, permitindo a visualização de todo território nacional em um formato A4.

O conhecimento da distribuição dos dados é vital para que se prossiga às etapas posteriores da execução. Para este tipo de análise podem ser usadas, por exemplo, tabelas com totais de ocorrência do fenômeno incluindo as variações temporais dos valores, se for o caso de estudo multitemporal como o apresentado aqui. Na Tabela 1 está apresentada a soma dos novos casos relatados por ano.

Tabela 1 - Número de casos de AIDS notificados por ano.

Fonte: MINISTÉRIO DA SAÚDE (2012).

ANO	CASOS	ANO	CASOS	ANO	CASOS	ANO	CASOS	ANO	CASOS
1980	1	1987	2847	1993	17396	1999	27735	2005	33166
1982	15	1988	4616	1994	19143	2000	30439	2006	32279
1983	41	1989	6326	1995	21881	2001	31064	2007	34128
1984	140	1990	9102	1996	24939	2002	35424	2008	36523
1985	554	1991	12110	1997	27358	2003	35424	2009	35980
1986	1193	1992	15181	1998	30273	2004	34193	2010	34217

A construção de mapas de pontos de contagem pode ser otimizada com base nos quatro princípios propostos por Arlinghaus (2005):

Princípio da Distribuição Aleatória: inicialmente, os pontos devem ser distribuídos aleatoriamente numa escala maior e então ser vistos em uma escala menor para visualização de padrões.

Princípio da Otimização: Consiste da aplicação do princípio anterior consecutivas vezes, usando uma hierarquia de camadas (*layers*) que se agrupam sucessivamente (como no exemplo brasileiro, municípios que formam unidades da federação, e estes por sua vez formam o país). Quando os dois primeiros princípios são seguidos, Arlinghaus considera um mapa de pontos de contagem propriamente construído e otimizado.

Princípio da Escala: quando a mudança em escala produz uma aglomeração de pontos, deve-se alterar a quantidade representada por cada ponto para manter a construção apropriada do mapa (ver item 2.2, abaixo).

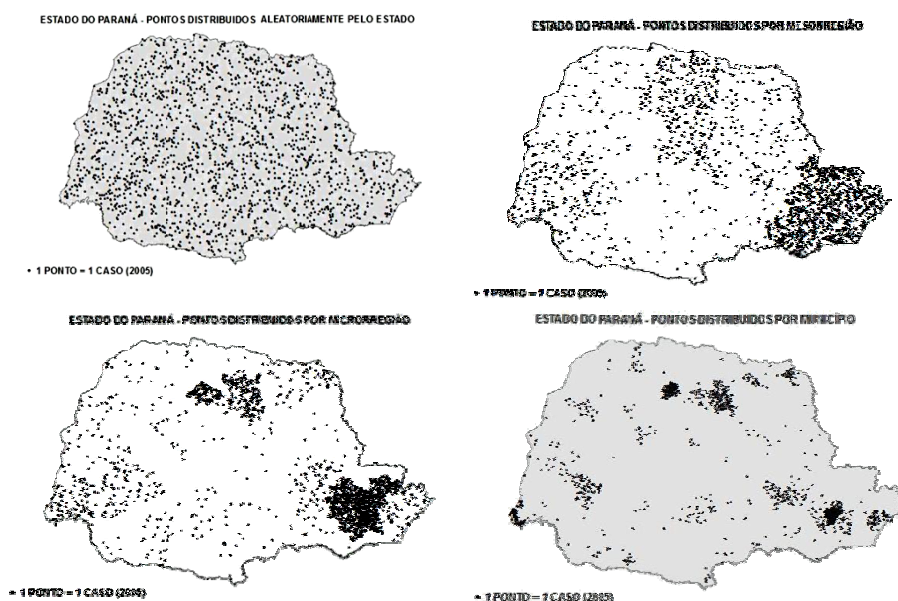
Princípio da Projeção: um mapa de pontos de contagem apropriadamente construído deve ser baseado numa projeção equivalente, uma vez que a quantidade de pontos por área (densidade de pontos) é a principal informação a ser transmitida pelo mapa.

A divisão do País em regiões homogêneas data de 1969, com a publicação pelo IBGE (Instituto Brasileiro de Geografia e Estatística) do estudo Divisão do Brasil em Microrregiões Homogêneas (LIMA et al, 2002). Neste estudo foi lançado o conceito de Espaço Homogêneo como sendo a “*forma de organização em torno da produção, expresso por combinações de fatos físicos, sociais e econômicos, permitindo a individualização de áreas*”. Portanto, o agrupamento dos municípios em regiões e destes em micro e mesorregiões, forma unidades consistentes entre si escolhidas por diversos fatores, e esta hierarquia pode ser indicada para aplicar o princípio da otimização descrito anteriormente.

A Figura 2 traz um exemplo da importância da escolha da unidade para distribuição aleatória: os pontos aleatoriamente distribuídos provocam uma visualização equivocada do fenômeno no primeiro mapa. Nos mapas seguintes, a

real distribuição do fenômeno fica mais clara à medida que são utilizadas subdivisões menores, até que o último mapa, que distribui os pontos no município de ocorrência, proporciona a visualização da distribuição espacial dos casos de AIDS no Paraná.

Figura 2 - Exemplo – Casos de AIDS no estado do Paraná em 2005, posicionados considerando diversas unidades de enumeração.



## 2.2 Determinação do número de ocorrências por ponto (N)

Em um mapa de pontos de contagem, quando é escolhido um valor de N muito pequeno (criando um excesso de pontos), ocorrem dois problemas: os pontos podem superpor-se criando áreas muito densas, e quando isolados os pontos podem causar uma falsa impressão de acurácia das informações.

Os pontos isolados transmitem a idéia de que o fenômeno ocorreu precisamente naquele local, e não que representa uma posição aleatória dentro de um polígono. Por outro lado, um valor de N muito grande pode representar o fenômeno retratado com um número muito pequeno de pontos, prejudicando a expressão da principal característica do mapa de pontos de contagem, que é a de demonstração de padrões de densidade espacial. Adicionalmente, pode

sobregeneralizar áreas com poucas ocorrências. Portanto, a escolha de N é um dos fatores críticos no projeto de um mapa por pontos de contagem.

O ideal, segundo Dent (1999), é que o polígono com o menor número de ocorrência seja representado por dois ou três pontos, e assim os demais valores são ajustados.

No caso em estudo, o estado com menos casos de AIDS é o Acre. Portanto, o valor de N para representar o Acre com pelo menos dois pontos nos anos finais da série é 20. A série histórica variando de 1 caso (1980) a 36.528 casos (número máximo de casos, alcançado em 2008, para depois diminuir nos anos subsequentes) impede que um valor de N único seja utilizado.

Por exemplo, se utilizasse  $N=20$  para toda a série, nos mapas dos anos iniciais da série temporal, o Brasil todo poderia ser representado por pouquíssimos pontos, sem representar a densidade espacial dos casos. Por outro lado, se fosse aplicado  $N=1$  para toda a série, o estado de São Paulo com 12.928 casos, em 1998, seria representado por uma nuvem indecifrável de pontos.

Portanto, foi escolhido um N variável por período: de 1980 até 1986,  $N=1$  caso; de 1987 a 1996,  $N=10$  casos; e de 1997 até 2010,  $N=20$  casos; conforme Tabela 02, com o total de casos notificados por ano. Por volta de 1997, há uma redução da taxa de crescimento das notificações, permitindo a continuidade do uso do valor de N igual a 20 casos.

No entanto, a escolha de valores diferentes de N para cada período pode dificultar a interpretação da evolução temporal do fenômeno pelo usuário, uma vez que 10 pontos representados individualmente em 1986 são todos agregados em um só ponto representado em 1987 (Figura 3). A primeira medida para evitar interpretações equivocadas é deixar em evidência o valor do N no mapa final, para que a mudança de agrupamento seja facilmente percebida. Existe também a possibilidade de se variar o tamanho dos pontos com o aumento do N. Esta opção leva ao aumento de complexidade da decisão do tamanho do ponto, que por sua vez é dependente de fatores como a sobreposição, o que pode prejudicar a correta interpretação do mapa.

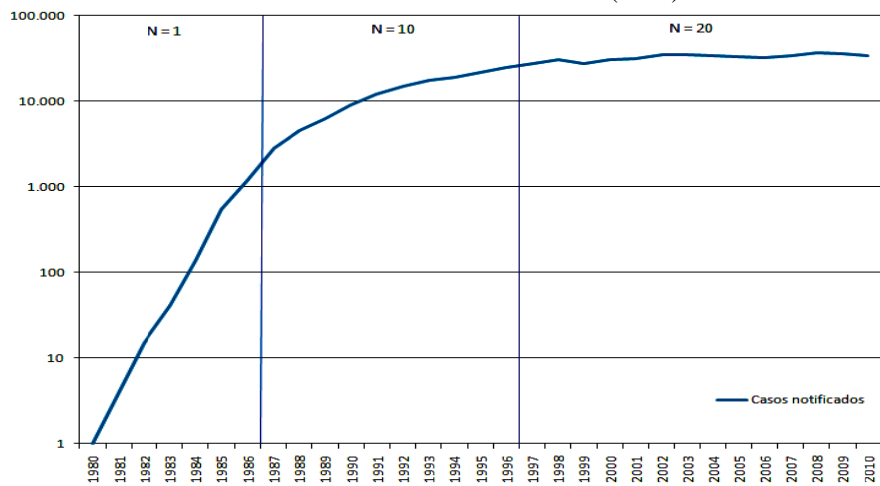
Tabela 2 - Total dos casos de AIDS notificados por ano, por estado.

Fonte: MINISTÉRIO DA SAÚDE

	80	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00	01	02	03	04	05	06	07	08	09	10	
RO	Q	Q	Q	Q	Q	Q	5	5	2	12	4	29	30	49	65	57	73	91	75	137	139	194	197	193	206	257	270	253	250	260	
AC	Q	Q	Q	Q	Q	Q	3	3	4	3	8	6	8	12	1	6	13	19	30	25	31	43	37	50	51	33	55	53	36	53	
AM	Q	Q	Q	Q	Q	Q	1	5	10	16	25	35	55	56	86	95	115	188	223	293	302	341	415	376	518	563	611	675	860	1053	1077
RR	Q	Q	Q	Q	Q	Q	1	1	7	6	13	8	6	8	8	17	15	23	24	50	54	34	81	104	94	70	108	143	155	161	
PA	Q	Q	Q	Q	1	3	9	21	31	35	64	88	116	154	182	228	246	295	295	297	512	573	629	917	844	851	1087	1392	1597	1476	
AP	Q	Q	Q	Q	Q	Q	Q	3	2	1	5	7	1	6	17	21	33	25	42	35	48	66	46	71	73	83	104	102	121	115	
TO	Q	Q	Q	Q	Q	Q	Q	Q	1	Q	9	10	7	15	22	27	25	34	53	56	57	68	78	79	89	74	94	89	126	153	132
MA	Q	Q	Q	Q	Q	Q	2	9	12	30	59	71	83	112	127	149	192	214	295	275	332	356	395	514	619	590	693	698	859	973	928
PI	Q	Q	Q	Q	Q	Q	2	8	15	12	20	33	28	29	66	78	82	102	97	111	174	161	190	186	256	268	236	325	331	364	344
CE	Q	Q	1	Q	5	12	18	47	66	92	174	231	217	267	340	338	373	647	570	561	597	650	797	848	743	746	1094	1090	1084	942	
RN	Q	Q	Q	1	1	Q	4	18	13	34	29	56	42	79	92	74	111	137	185	134	171	170	157	207	143	263	229	333	332	352	335
PB	Q	Q	Q	Q	Q	7	4	23	18	20	54	68	74	104	124	124	134	151	195	252	243	236	265	281	288	288	265	301	338	348	395
PE	Q	Q	Q	1	1	8	16	68	90	167	182	227	258	316	356	397	569	593	776	678	757	916	1216	1117	1136	1329	1235	1258	1314	1206	1500
AL	Q	Q	Q	Q	Q	Q	8	8	29	21	39	43	70	82	83	90	123	154	119	134	154	194	203	234	202	265	247	354	454	343	330
SE	Q	Q	Q	Q	Q	Q	Q	11	20	19	22	30	49	55	92	85	81	104	104	114	116	93	112	154	146	162	180	226	275	223	245
BA	Q	Q	Q	Q	2	5	27	60	101	175	187	278	344	381	342	358	521	619	602	721	879	910	1157	1167	1233	1216	1176	1361	1515	1662	1682
MG	Q	1	1	6	41	65	135	194	286	458	608	778	1286	1545	1654	1995	1829	1800	1837	1812	2084	2559	2725	2652	2955	2504	2690	2680	2468	2344	
ES	Q	Q	Q	Q	4	8	19	33	51	58	85	114	165	217	211	251	369	410	433	490	547	642	651	669	622	613	568	645	682	717	
RJ	Q	4	7	27	112	295	605	910	1182	1456	1682	2273	2359	2509	2898	3579	4054	4204	3753	4277	4268	4980	4682	5036	4715	4508	4437	4309	4931	4504	
SP	1	9	27	88	327	597	1548	2573	3476	5154	8710	8303	8958	9328	10555	11485	11667	12753	11058	11162	10570	11162	10774	9322	8350	8148	7783	8016	7477	8577	
PR	Q	1	Q	1	6	16	39	76	111	180	317	463	596	694	915	1162	1455	1557	1611	1841	1768	2047	2024	1869	1300	1710	1969	3028	2096	1984	
SC	Q	Q	Q	Q	1	1	10	24	57	90	231	332	434	570	748	972	1143	1187	1505	1393	1660	1808	2081	1961	1928	1371	1895	1700	1807	1932	1886
RS	Q	Q	Q	3	7	14	85	123	229	323	466	687	724	996	1186	1391	1597	2161	2854	2543	3292	3459	4055	4148	3623	3485	3823	4567	4446	4232	4018
MS	Q	Q	Q	2	3	11	13	25	34	91	135	152	269	243	243	282	289	298	265	279	290	394	419	402	369	402	425	499	451	438	
MT	Q	Q	Q	2	4	8	19	18	25	47	77	101	104	174	200	298	343	300	204	273	380	521	578	609	580	491	594	560	570	527	
GO	Q	Q	Q	Q	2	5	9	52	71	70	100	152	233	266	366	480	400	578	503	489	647	730	773	772	825	775	780	774	684	787	840
DF	Q	Q	Q	Q	Q	5	10	23	41	72	86	206	227	220	249	272	327	377	340	344	416	334	462	588	455	445	399	383	412	434	406
	N = 1						N = 10												N = 20												

Figura 3 - Evolução dos Casos Notificados e N escolhido

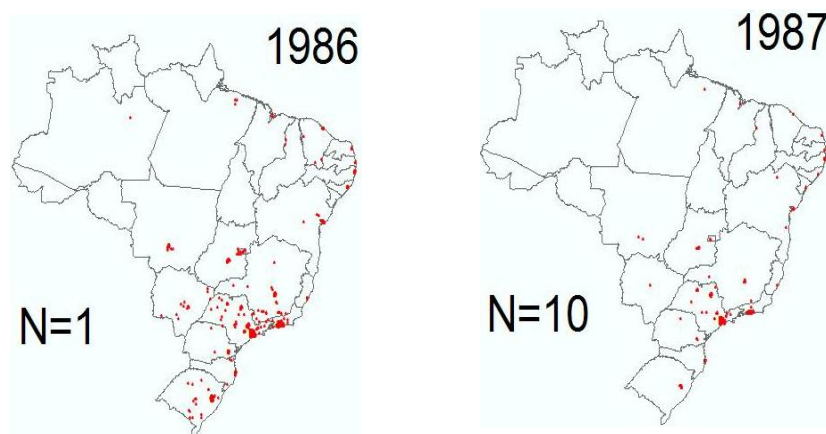
Fonte: MINISTÉRIO DA SAÚDE (2012)





A Figura 4 ilustra a questão da troca do valor de N ao longo da série temporal. O número de pontos no mapa de 1987 é menor, apesar dos casos terem aumentado de 1193 para 2847, porque neste mapa cada ponto representa 10 casos, ao contrário de apenas 1 no mapa anterior. Existem na literatura outras situações similares de conflitos na definição de N para séries temporais, como o retratado em CDC (2008), onde o N foi definido como 50 casos de AIDS notificados nos Estados Unidos em 2005. Como os casos de AIDS nos EUA em 2008, data do trabalho, eram muito numerosos (975.350 - os mapas representam valores cumulativos), CDC (2008) consideraram a alteração de N para 100, mas não o fizeram com receio que esta mudança pudesse subestimar a percepção dos casos em muitas regiões.

Figura 4 - Efeito de diferentes valores de N na distribuição dos casos de AIDS no Brasil em 1986 e 1987.



### 2.3 Tamanho do Ponto

Na decisão a respeito do tamanho do símbolo utilizado para os pontos, considera-se que na área mais densa os pontos devem se fundir levemente (o termo em inglês usado é *coalescence*). Tradicionalmente, é utilizado para mapas de pontos de contagem analógicos o Nomógrafo de Mackay, de 1949 (Figura 5), que ajuda o engenheiro cartógrafo a ajustar o tamanho do ponto e também o valor de N com base na densidade espacial do fenômeno, por centímetro quadrado, na zona em que os pontos começam a se fundir (a área mais densa do mapa).

Kimerling (2008) realizou uma atualização desses estudos, criando uma formulação matemática que permitisse o cálculo da sobreposição entre os pontos. Como resultado, obteve graduações de diferentes porcentagens de cobertura para tamanhos de pontos variados (Figura 6). Neste trabalho, considerando a escala final de visualização dos mapas de 1:30.000.000, a densidade de pontos por cm<sup>2</sup> por

Unidade da Federação (UF) pode ser observado na Tabela 3. Portanto, segundo as graduações propostas por Kimerling (2008), o ideal é que se usasse o tamanho de 1 pt, para que as UF com maior densidade de pontos, Rio de Janeiro e Distrito Federal, tivessem uma sobreposição de 10% a 30%. No entanto, como pode ser observado na Figura 7, nas áreas com poucos casos, o tamanho do ponto se tornava muito pequeno prejudicando a interpretação do mapa. Portanto, foi definido o tamanho de 2 pt para os pontos, de forma que apenas as duas UF citadas tivessem sobreposição maior que 50%, e as demais permanecessem com a visualização de casos adequada.

Figura 5 - Nomógrafo de Mackay, 1949.

Fonte: KIMERLING (2008).

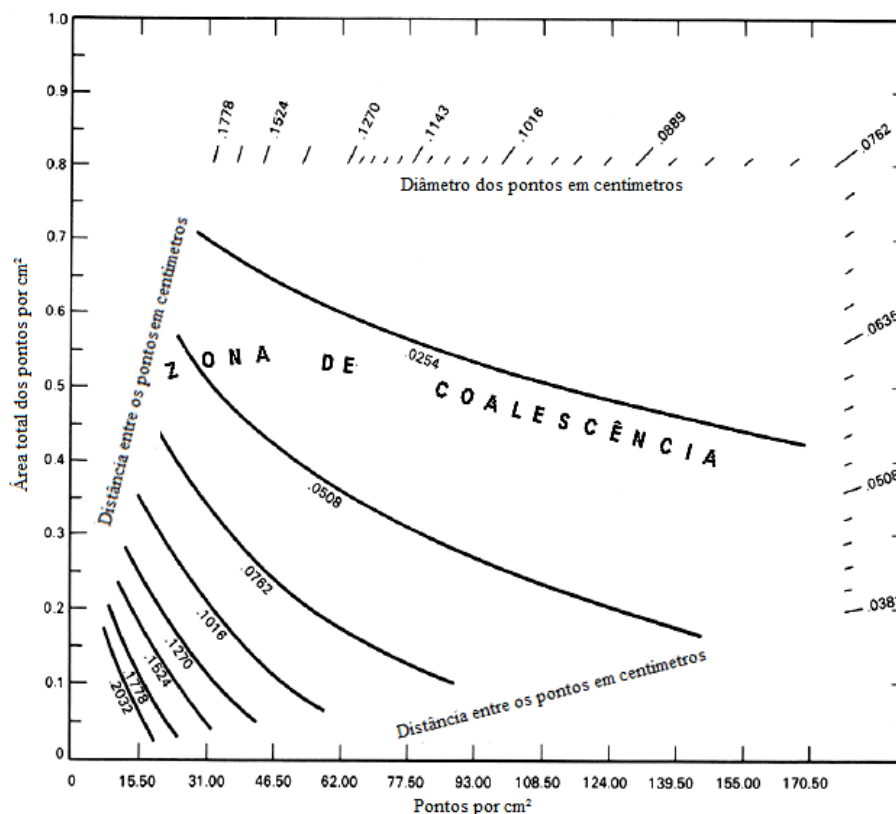


Figura 6 - Variação da sobreposição para vários tamanhos de pontos

Fonte: KIMERLING (2008).

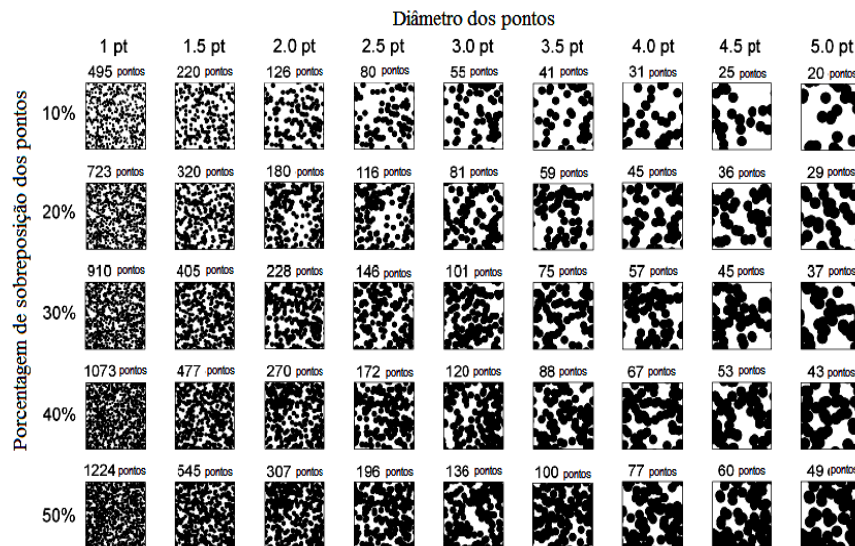


Figura 7 - Casos em 2010 representados com diferentes tamanhos de ponto.

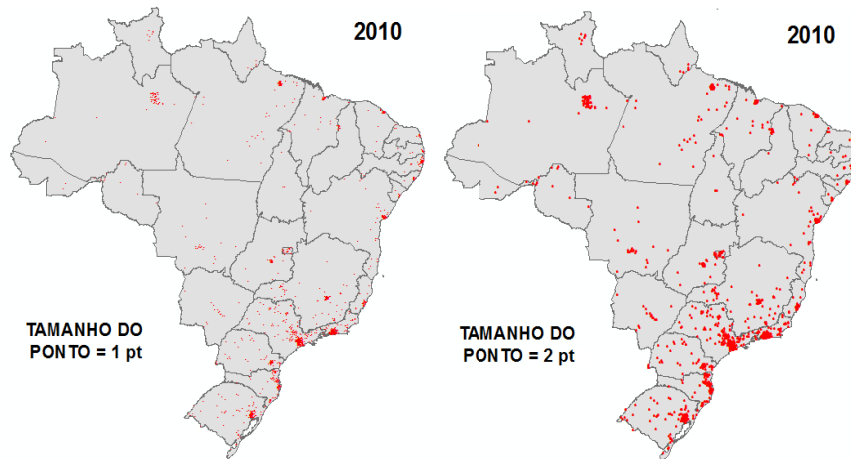


Tabela 3 - Densidade máxima de casos por UF.

UF	Área em km <sup>2</sup> (IBGE)	Área em cm <sup>2</sup> (escala 1:30.000.000)	Número máximo de Pontos	Máximo de pontos por cm <sup>2</sup>
RO	237.576,17	2,64	27	10
AC	152.581,39	1,70	6	3
AM	1.570.745,68	17,45	108	4
RR	224.298,98	2,49	16	4
PA	1.247.689,52	13,86	160	8
AP	142.814,59	1,59	12	7
TO	277.620,91	3,08	15	3
MA	331.983,29	3,69	97	19
PI	251.529,19	2,79	36	12
CE	148.825,60	1,65	109	66
RN	52.796,79	0,59	35	57
PB	56.439,84	0,63	40	48
PE	98.311,62	1,09	150	115
AL	27.767,66	0,31	45	115
SE	21.910,35	0,24	28	93
BA	564.692,67	6,27	168	22
MG	586.528,29	6,52	339	40
ES	46.077,52	0,51	72	111
RJ	43.696,05	0,49	716	914
SP	248.209,43	2,76	2297	282
PR	199.314,85	2,21	303	89
SC	95.346,18	1,06	229	160
RS	281.748,54	3,13	457	146
MS	357.124,96	3,97	56	11
MT	903.357,91	10,04	61	6
GO	340.086,70	3,78	96	20
DF	5.801,93	0,06	65	594

Fonte das áreas: IBGE (2002).

## 2.4 Posicionamento dos pontos

Uma vez determinada a unidade de enumeração, o tamanho dos pontos e o valor de N, a próxima etapa consiste do posicionamento destes pontos no mapa. Slocum (2009) descreve quatro diferentes técnicas para o posicionamento dos pontos na área de ocorrência. A primeira é a técnica uniforme, que considera os pontos distribuídos uniformemente na área de enumeração. Quando os fenômenos não ocorrem distribuídos uniformemente esta técnica de posicionamento não é adequada. A segunda técnica é a randômica, que usa uma distribuição aleatória dos

pontos, como a usada no primeiro exemplo da Figura 2. Esta técnica pode revelar *clusters* e padrões inexistentes, gerados pela aleatoriedade e não pela característica de distribuição do fenômeno. A técnica de Pesos Geográficos usa a auto-correlação espacial para detectar *clusters*, mas é necessária a informação adicional da ocorrência pontual do fenômeno, o que nem sempre está disponível. Finalmente, há a técnica geograficamente embasada, onde um dado auxiliar mais detalhado está disponível e é usado para auxiliar o posicionamento. Estes dados podem ser tanto divisões administrativas menos globais que as áreas de enumeração, como é o caso deste trabalho, como máscaras que apontem regiões com maior probabilidade de ocorrência do fenômeno.

## 2.5 Legenda

Tomadas as decisões principais do projeto do mapa de pontos de contagem, a etapa final consiste em comunicar de forma eficaz ao usuário as decisões tomadas nas etapas anteriores através da construção da legenda. Para tanto, é importante ressaltar que o usuário pode ter dificuldade em estimar corretamente a densidade de pontos em determinada região do mapa, sendo possível usar valores âncoras para ajudar a comparabilidade numérica do fenômeno, se esta habilidade for importante para atingir o objetivo definido para o mapa.

## 3. DESENVOLVIMENTO E RESULTADOS

### 3.1 Agrupamento dos Pontos

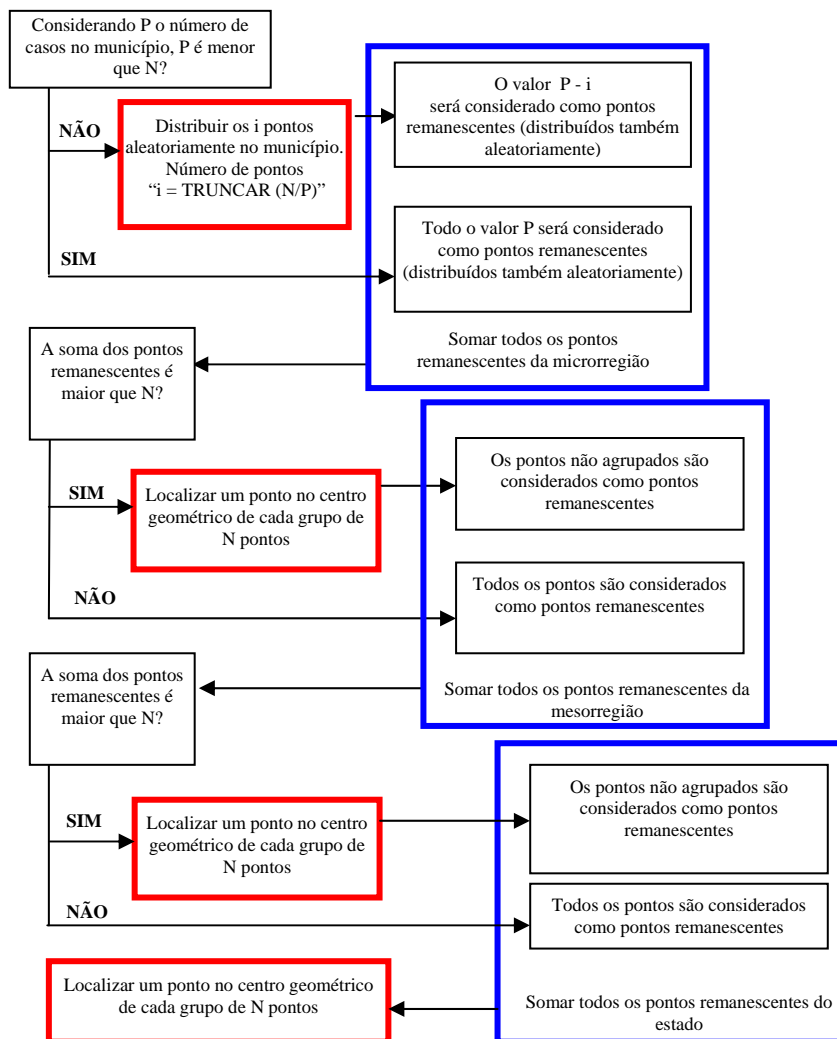
Como estabelecido anteriormente, a unidade de enumeração dos mapas deste estudo foi a Unidade da Federação. No entanto, para a distribuição dos pontos pelo método *Geograficamente Embasado* foram usados os polígonos dos limites municipais, que permitem distribuir os pontos de acordo com sua probabilidade de ocorrência.

Como o crescimento na ocorrência de casos de AIDS se deu de forma gradual pelo País, no começo da série histórica existiam muitos municípios com o número de casos menor que N. Mesmo em regiões com mais casos, são representados apenas o número de pontos múltiplo de N. Nestas situações é gerado um excedente de pontos não representados que se denomina pontos remanescentes. O principal desafio deste trabalho foi a automatização do agrupamento dos pontos remanescentes para que não sejam desconsiderados casos.

A questão dos agrupamentos dos pontos é recorrente na simplificação de mapas de pontos de contagem. Por exemplo, Berg et al. (2002), avaliou diversas heurísticas para agrupamento de pontos, em busca da obtenção de valores que representassem o quão bem o conjunto original de pontos se encontrava representado pelos pontos agregados. Neste trabalho esta questão também foi central na definição da metodologia, sendo utilizado o conceito hierárquico de

Divisão do Território Brasileiro em microrregiões, mesorregiões e Unidades da Federação para sucessivamente agrupar os pontos remanescentes conforme digrama mostrado na Figura 8.

Figura08 - Diagrama do algoritmo para agrupamento de pontos.

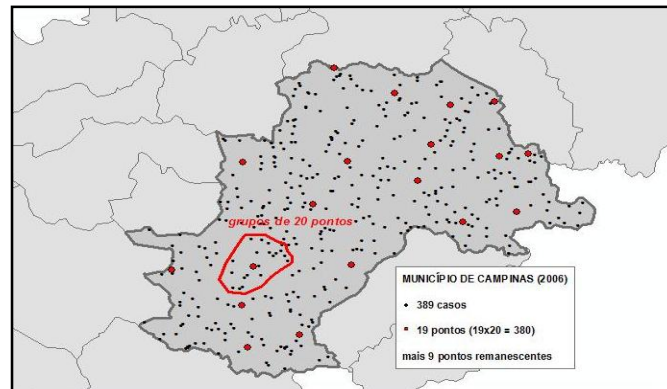


O algoritmo apresentado na Figura 8 foi implementado usando a Ferramenta *Model Builder*, do software ArcGIS 9.2, para que os agrupamentos se dessem de forma automática para cada ano. Abaixo na Figura 9, a mesorregião de Campinas

ilustra a aplicação do método. Esta mesorregião é uma das quinze do estado de São Paulo e é formada pela união de 49 municípios agrupados em cinco microrregiões. No exemplo, foram usados os casos de 2006, que foram 694 para a mesorregião, e foi adotado  $N = 20$ .

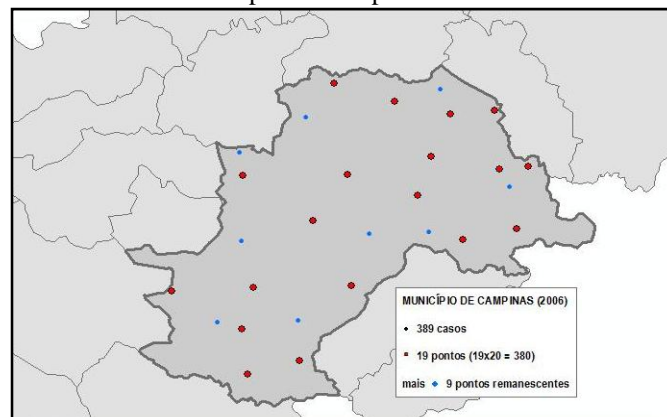
- Inicia-se a aplicação do método com a observação de cada município, que é a unidade de enumeração inicial. Os 389 casos do município de Campinas, por exemplo, são representados por 19 pontos aleatoriamente distribuídos no município.

Figura 9 - Distribuição inicial dos casos de AIDS no Município de Campinas (2006).



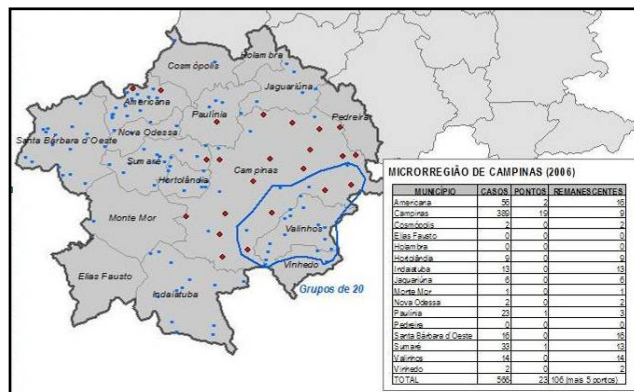
- Com os 19 pontos representaram-se 380 dos casos, gerando um excesso de 9 casos remanescentes, apresentados em azul na Figura 10.

Figura 10 - Pontos no município de Campinas e casos remanescentes.



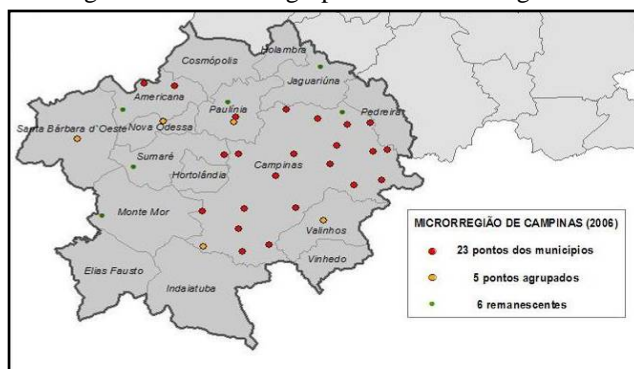
- Repetindo o procedimento para todos os municípios da microrregião, chegou-se a 23 pontos (representados em vermelho) e 106 casos remanescentes apresentados em azul. São 106 casos remanescentes que geram cinco pontos adicionais.

Figura 11 - Agrupamento dos pontos e casos remanescentes.



- Com os 5 pontos gerados dos casos remanescentes por município agrupados, ainda restam 6 remanescentes da microrregião (100 casos representados por 20 pontos)

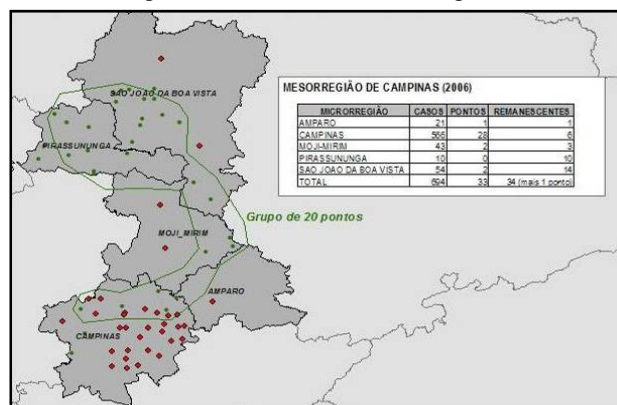
Figura 12 - Pontos Agrupados na Microrregião.



- Executando o procedimento para todas as microrregiões da mesorregião de Campinas, foi gerado um excedente de 34 casos, que agrupados se transformam em 1 ponto adicional.



Figura 13 - Pontos na Mesorregião.



Foram criados 30 mapas da série anual para todo o Brasil, dos quais foram selecionados sete para compor a Figura 14:

#### 4. CONCLUSÃO

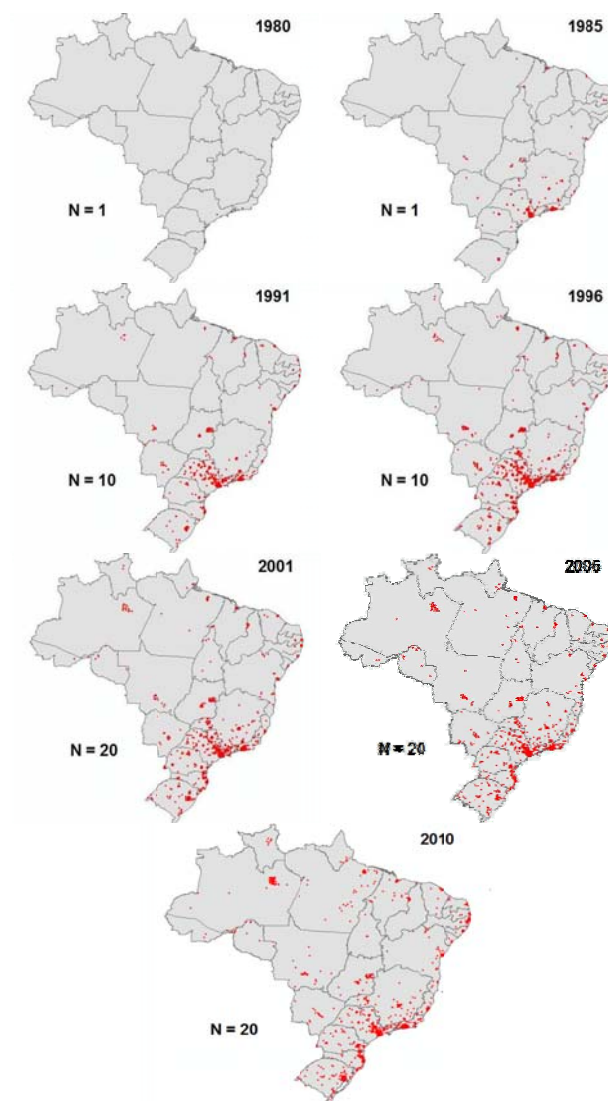
Os mapas de pontos de contagem são ferramentas eficientes, pois possibilitam uma visualização clara do crescimento da doença no país ao longo do tempo.

Utilizar a distribuição do espaço brasileiro estabelecida pelo IBGE permite automatizar a distribuição dos pontos em regiões hierarquizadas, o que refina o resultado sem o trabalho intenso do agrupamento manual dos pontos. A mesma sistemática pode ser utilizada para agrupar dados de setores censitários em distritos e posteriormente em municípios, dependendo dos dados existentes e do propósito do mapa. Uma vez que os dados históricos dos recenseamentos no Brasil se encontram georreferenciados pela base de setores censitários, a metodologia tem potencial para geração de mapas temáticos para uma gama extensa de variáveis.

Alguns questionamentos permanecem, como, por exemplo, sobre a melhor forma de representar uma série histórica com diferentes valores de N. Uma possibilidade é a adição de uma nova variável visual, o tamanho ou tom de cor do símbolo, para representar valores distintos do fenômeno através do tempo. No entanto, testes com usuários são necessários para esclarecer o êxito desta técnica em tornar mais claro o entendimento da distribuição da variável.

O próprio agrupamento dos pontos pode ser ainda otimizado, utilizando-se técnicas adicionais de agregação de dados e identificação de *clusters* (agrupamentos). A melhoria da distribuição dos pontos pode ser obtida com a inserção de um mapa de restrições, baseado em fatores como uso do solo, por exemplo, que permitisse apenas pontos nas áreas urbanas e até uma determinada distância das rodovias, por exemplo, ou segundo outros critérios específicos de influência na propagação dos casos.

Figura 14 - Sete dos mapas gerados com a distribuição anual de novos casos de AIDS no Brasil.



## AGRADECIMENTOS

Agradecemos aos colegas Gisele Nepomuceno Canova e Marcio Augusto Reolon Schmidt pelas valiosas contribuições a este artigo.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ARLINGHAUS, Sandra L, 2005. "Spatial Synthesis: The Evidence of Cartographic Example: Hierarchy and Centrality." *Solstice: An Electronic Journal of Geography and Mathematics*, Volume XVI, Number 1. Ann Arbor: Institute of Mathematical Geography, 2005.
- CDC, 2008 "*AIDS Dot Maps (through 2005)*" Centers for Disease Control and Prevention, Department of Health and Human Services, EUA. Disponível em: <http://www.cdc.gov/hiv/topics/surveillance/resources/slides/dot/index.htm> (acessado em 22/05/2012)
- DENT, Borden D., 1999 "*Cartography - Thematic Map Design*" WCB/McGraw-Hill, 5ª edição.
- DSHS, 2008 "*Texas AIDS cases 1986 – 2007*" Department of State Health Services HIV/STD Program. Disponível em: <http://www.dshs.state.tx.us/dotmaps.shtm> (acessado em 22/05/2012).
- ENG, Steven B.; WERKER, Denise H.; KING, Arlene; MARION, Stephen A.; BELL, Alison; ISSAC-RENTON, Judith L.; IRWIN, G. Stewart e BOWIE, WILLIAM R., 1999 "Computer-Generated Dot Maps as an Epidemiologic Tool: Investigating an Outbreak of Toxoplasmosis". Em: *Emerging Infectious Diseases Journal* Vol. 5, No. 6 Nov–Dec.
- JIANG, Hong, 2004 *Notas de Aula – Disciplina Introdução à Cartografia – The University of Wisconsin – Madison Department of Geography*. Disponível em: [http://www.geography.wisc.edu/classes/geog370/labs/5\\_dot\\_maps\\_s04.pdf](http://www.geography.wisc.edu/classes/geog370/labs/5_dot_maps_s04.pdf) (acessado em 22/05/2012).
- KIMERLING, A.J., 2008 "*Dotting the Dot Map, Revisted*" University of Redlands Masters of Science in GIS Program Colloquium Series Redlands, California Disponível em: [http://downloads2.esri.com/MappingCenter2007/resources/presentations/Kimerling\\_2008\\_UR\\_Colloquium.pdf](http://downloads2.esri.com/MappingCenter2007/resources/presentations/Kimerling_2008_UR_Colloquium.pdf) (acessado em 22/05/2012).
- LIMA, Maria Helena Palmer; RODRIGUES, Cristiane Moreira; SILVA, Jorge Kleber Teixeira; MARTINS, Paulo Cesar; TERRON, Sônia Luiza; SILVA, Rosângela Lemos de Souza, 2002 "*Divisão Territorial Brasileira*" Fundação Instituto Brasileiro de Geografia e Estatística. Disponível em: [ftp://ftp.ibge.gov.br/Censos/Censo\\_Demografico\\_2000/Anexos/CD\\_1/Desenho.zip](ftp://ftp.ibge.gov.br/Censos/Censo_Demografico_2000/Anexos/CD_1/Desenho.zip) (acessado em 22/05/2012).
- MINISTÉRIO DA SAÚDE, 2012 "*Banco de Dados DST – AIDS*" Programa Nacional de DST e AIDS. Disponível em: [http://www2.aids.gov.br/final/dados/dados\\_aids.asp](http://www2.aids.gov.br/final/dados/dados_aids.asp) (acessado em 22/05/2012).

SLOCUM, Terry A.;McMASTER, Robert B; KESSLER, Fritz C e HOWARD, Hugh H, 2009 “*Thematic Cartography and Geovisualization*” Prentice Hall; 3ª edição.

(Recebido em junho de 2012. Aceito em dezembro de 2012.)