

INVESTIGAÇÃO DE AGLOMERADOS ESPACIAS DE DOENÇAS EM APIÁRIOS NO ESTADO DO RIO DE JANEIRO UTILIZANDO PROGRAMAÇÃO INTEIRA

INVESTIGATION OF PARTICLE SPATIAL DISEASES IN BEEKEEPING RIO DE JANEIRO USING FULL SCHEDULE.

Larissa M. De Souza¹, Marcelo Dib Cruz¹, Maria Cristina Afonso Lorezon², Wagner S. Tassinari³

¹ Curso de Engenharia de Agrimensura e Cartográfica, Universidade Federal Rural do Rio de Janeiro (UFRRJ), Rodovia BR-465, km 7, CEP 23890000 - Seropédica/RJ. E-mail: larissamessiass@gmail.com

^{1'} Departamento de Matemática, Universidade Federal Rural do Rio de Janeiro (UFRRJ), Rodovia BR-465, km 7, CEP 23890000 - Seropédica/RJ. E-mail: madibcruz@gmail.com.

² Zootecnista, D.Sc. Departamento de Produção Animal, Instituto de Zootecnia, Universidade Federal Rural do Rio de Janeiro (UFRRJ), Rodovia BR-465, km 7, CEP 23890000 - Seropédica/RJ. E-mail: lorenzon_ufrj@yahoo.com.br

³Epidemiologista e Estatístico, D.Sc. Departamento de Matemática e Curso de Pós-Graduação em Ciências Veterinárias, Instituto de Ciências Exatas, Universidade Federal Rural do Rio de Janeiro (UFRRJ), Rodovia BR-465, km 7, CEP 23890000 - Seropédica/RJ. E-mail: tassinari@ufrj.br

ABSTRACT. The study aims to identify possible areas with the highest concentration of bee diseases and disease outbreaks in the state of Rio de Janeiro. For this, the problem will be studied, modeled and will be built using some heuristics Metaheuristics Genetic Algorithms, GRASP and ILS. The heuristics will be compared to see which of metaheuristics best fits the problem.

INDEX TERMS: Space clusters, diseases, beekeeping, Metaheuristics.

RESUMO. O presente estudo objetiva identificar possíveis áreas com maior concentração de doenças apícolas e os surtos de doenças no Estado do Rio de Janeiro. Para isto, o problema será estudado, modelado e serão construídas algumas heurísticas utilizando as Metaheurísticas Algoritmos Genéticos, GRASP e ILS. As heurísticas serão comparadas para verificar qual das metaheurísticas melhor se adapta ao problema.

1. INTRODUÇÃO

Nos últimos dez anos, colônias de abelhas começaram a desaparecer em diferentes países do mundo. Nos Estados Unidos, metade da população do inseto foi dizimada. Em algumas regiões da Europa, a diminuição chega a 80%. Existem hoje, no mundo, cerca de 25 mil espécies de abelhas. A importância delas vai além da produção de mel e própolis. Cerca de 70% da produção agrícola mundial depende destes insetos. Desta forma, o desaparecimento das colmeias tem impacto direto na produção de alimentos. Entre diversos fatores apontados como responsáveis pelo sumiço de abelhas, tais como o uso excessivo de pesticidas, como os neonicotinoides, poluição do ar e da água, além do estresse causado pelo gerenciamento inadequado das colmeias, o excesso de parasitas que afetam esses insetos está entre os principais.

A década do ano 2000 assinala, no Brasil e no mundo, o momento de exacerbação do perfil de mortalidade das abelhas por fatores diversos como o intenso comércio de colmeias e produtos apícolas. Sendo assim, um dos principais causadores da propagação das doenças das abelhas melíferas, que aliados à multiplicidade de outros fatores, exacerbam o quadro estrutural de sanidade, em déficit, pela falta da aplicação de medidas profiláticas, pelo despreparo técnico, pela ausência de padrões de seleção de colmeias, pelos cuidados com o meio ambiente, etc.

Na apicultura, as perdas anuais de colmeias, acima de 20%, traduz a relevância do diagnóstico e do controle de doenças e parasitoses. As autoridades responsáveis pela apicultura nacional estão continuamente confrontadas com novos quadros clínicos e agentes patogênicos. Conseqüentemente, há necessidade de se disponibilizar de exames laboratoriais para se obter diagnósticos das patologias apícolas como medidas essenciais de controle. Desde a década de 80, surge um renovado interesse nos estudos de padrões espaciais e temporais de doenças, conforme salienta a extensa literatura publicada em periódicos de diferentes áreas, incluindo importantes revisões (SMITH, 1982; WILLIAMS, 1984; GESLER, 1986; HILLS e ALEXANDER, 1989; MARSHALL, 1991; KNOX, 1991; RICHARDSON, 1992; WERNECK e STRUCHINER, 1997).

Dentre os desenhos epidemiológicos utilizados neste contexto, destacam-se os estudos de agregados (clusters, na língua inglesa) (KLEINBAUM et al., 1982). De maneira geral, aglomerados espaciais de doenças podem ser atribuídos aos fatores demográficos, genéticos, ambientais ou, socioculturais superpostos geograficamente ao padrão de ocorrência observado (MARSHALL, 1991). O estudo de técnicas para detecção de aglomerados espaciais no campo

da Epidemiologia recebe importante contribuição científica (WERNECK e STRUCHINER, 1997).

O problema acima descrito pode ser definido na área de Otimização como um problema de Clusterização ou Agrupamento. Clusterização é o termo genérico para um processo que une objetos similares em um mesmo grupo. Cada grupo é denominado um cluster. O número de clusters pode ser conhecido ou não. Quando o número de clusters é conhecido, tratamos Problema de K-Clusterização ou, simplesmente Problema de Clustrização (PC). Caso contrário é denominado Problema de Clusterização Automática (PCA). A utilização de Programação Linear Inteira permite que certos problemas sejam modelados (ou formulados) e resolvidos de maneira exata por meio da utilização de pacotes de software.

Na busca por respostas de caráter epidemiológico, faz-se necessário conhecer quais fatores ambientais e de manejo mais influenciam a incidência de doenças apícolas no estado do Rio de Janeiro, visando uma busca de medidas e soluções que reduzam sua incidência e, conseqüentemente, permitam uma melhoria no quadro produtivo dos apiários. Para isso é necessário estudar as regiões onde encontram as maiores incidências de doenças. Este trabalho objetiva utilizar modelos exatos de Programação Inteira para a investigação da existência de aglomerados espaciais de doenças em apiários.

1. MATERIAL E MÉTODOS.

2.1 Áreas de Estudo

O presente estudo foi realizado em 38 municípios do estado do Rio de Janeiro que está situado a 22°54'23" de latitude sul e 43°10'21" de longitude oeste, na porção leste da região Sudeste do Brasil, tendo como limites os estados de Minas Gerais (norte e noroeste), Espírito Santo (nordeste) e São Paulo (sudoeste), como também o Oceano Atlântico (leste e sul) (FIBGE, 2006). Os municípios que participaram do levantamento foram: Nova Friburgo, Itaocara, Paty do Alferes, Itaguaí, Paraíba do Sul, Teresópolis, São Fidélis, Sumidouro, Rio Claro, Sapucaia, Piraí, Resende, Trajano de Moraes, Barra do Piraí, Miguel Pereira, Petrópolis, Valença, Itatiaia, São José do Vale do Rio Preto, Tanguá, Araruama, Três Rios, Barra Mansa, Cambuci, Engenheiro Paulo de Frontin, Porciúncula, Rio Bonito, Vassouras, Casimiro de Abreu, Paracambi, Silva Jardim, Natividade, Aperibé, Comendador Levy Gasparian, Mendes, Pinheiral, Santa Maria Madalena e Volta Redonda, abrangendo todas as mesorregiões do estado do Rio de Janeiro.

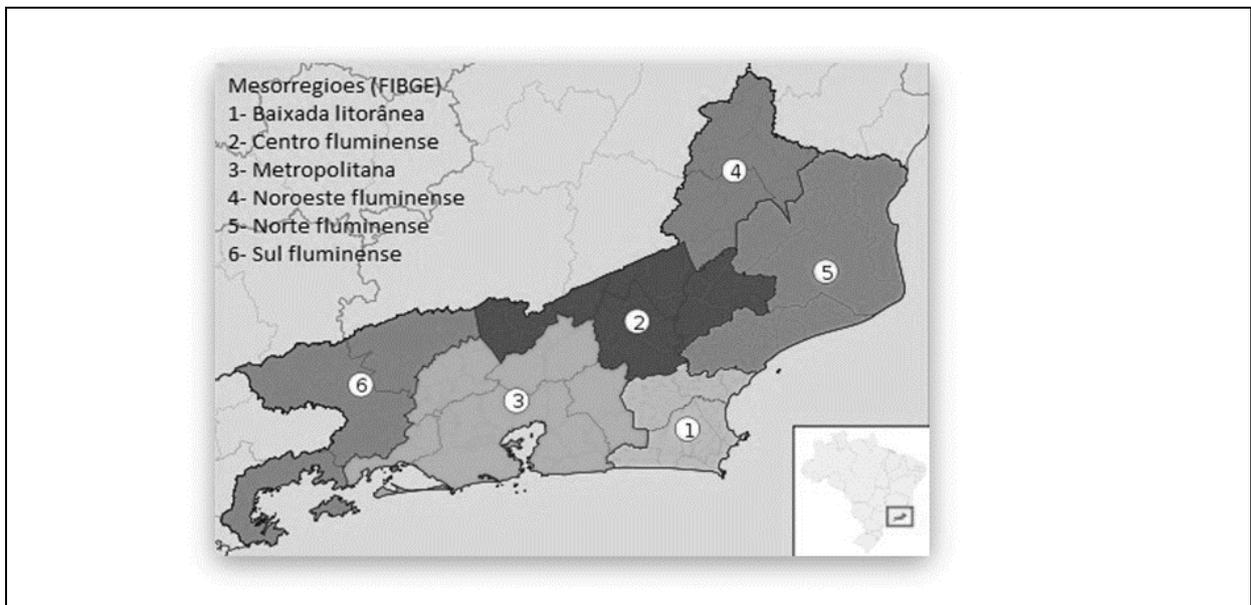


Figura 1: Distribuição Geográfica das Mesorregiões do estado do Rio de Janeiro.

2.2 Origens dos Dados

Os dados utilizados foram provenientes do levantamento dos déficits da apicultura do estado do Rio de Janeiro, e foram coletados no período de 2009 a 2010, por meio de aplicação de questionários por técnicos dos Núcleos de Defesa Agropecuária (NDA's) do Rio de Janeiro. Ao todo, 401 apicultores participaram deste levantamento, mas para o presente estudo foram selecionados aqueles que responderam sobre a ocorrência de doenças em seus apiários (n= 269). As coordenadas geográficas dos apiários foram identificadas, via GPS, no momento da aplicação dos questionários e utilizadas em formato UTM.

2.3 Metodologia

Para este trabalho, foram propostos dois modelos exatos de Programação Inteira para o problema de Clusterização. O primeiro proposto por Rao (1971) e o segundo proposto por Vinod (1969) e conhecido como k-mediana, conforme modelos descritos abaixo.

O primeiro modelo (RAO, 1971), é um modelo linear inteiro 0-1. O objetivo deste modelo é minimizar a maior distância interna de todos os clusters, ou ainda, minimizar a maior diagonal de todos os clusters. Cada ponto só pode pertencer a um cluster e o número de clusters k é um dado de entrada. A grande dificuldade deste modelo é que o número de restrições cresce rapidamente, à medida que as quantidades de pontos m e de clusters k crescem. Portanto, o modelo exato Diâmetro é indicado somente para estes valores pequenos. O modelo é descrito abaixo:

$$\text{Min } Z = D_{\max} \quad (16)$$

$$\text{S.A. } D_l \geq d_{ij} (x_{il} + x_{jl} - 1) \quad \forall i, j, l, \quad i = 1, \dots, m, \quad j = 1, \dots, m, \quad l = 1, \dots, k \quad (17)$$

$$\sum_{l=1}^k x_{il} = 1 \quad \forall i, \quad i = 1, \dots, m \quad (18)$$

$$D_{\max} \geq D_l \quad \forall l, \quad l = 1, \dots, k \quad (19)$$

$$x_{il} \in \{0, 1\} \quad \forall i, l, \quad i = 1, \dots, m, \quad l = 1, \dots, k$$

$$D_l \geq 0 \quad \forall l, \quad l = 1, \dots, k$$

onde $x_{ij} = \begin{cases} 1, & \text{se o ponto } i \text{ pertence ao cluster } j \\ 0, & \text{caso contrário.} \end{cases}$

d_{ij} = distância entre o ponto i e o ponto j .

D_l = diâmetro do cluster l .

D_{\max} = o maior diâmetro de todos os clusters.

A restrição (17) garante que a diagonal de cada cluster é maior ou igual a distância máxima entre dois pontos dentro do cluster. A restrição (18) garante que cada ponto só pode ser alocado a um cluster e, finalmente, a restrição (19) garante que D_{\max} seja maior que todas as outras diagonais de todos os clusters.

Já o segundo modelo (VINOD, 1969), o objetivo é selecionar k pontos médios do conjunto de pontos X , um ponto médio para cada cluster, onde a soma das distâncias dos outros pontos do cluster a este ponto médio é mínima. Cada ponto só pode pertencer a um cluster e o número de clusters k é um dado de entrada. O ponto médio de um cluster é definido como sendo aquele ponto que pertence ao cluster e que está mais próximo a todos os outros pontos do cluster.

O modelo K-medianas é mostrado abaixo. A restrição (13) assegura que cada ponto será associado a somente um ponto médio. Restrição (14) assegura que serão escolhidos exatamente k pontos médios e a restrição (15) que cada ponto i está associado a um ponto j se e somente se o ponto j é um ponto médio.

$$\text{Min } \sum_{i=1}^n \sum_{j=1}^n d_{ij} x_{ij}$$

(1)

$$\text{S.A. } \sum_{j=1}^n x_{ij} = 1 \quad \forall i, \quad i = 1, \dots, n$$

(2)

$$\sum_{j=1}^n x_{jj} = k$$

(3)

$$x_{ij} \leq x_{jj} \quad \forall i, j, \quad i = 1, \dots, n \quad j = 1, \dots, n$$

(4)

$$x_{ij} \in \{0,1\} \quad \forall i, j, \quad i = 1, \dots, n, \quad j = 1, \dots, n$$

(5)

onde $x_{ij} = \begin{cases} 1, & \text{se o ponto } i \text{ está associado ao ponto médio } j \\ 0, & \text{caso contrário.} \end{cases}$

$x_{jj} = 1$, significa que o ponto j é o ponto médio.

d_{ij} = distância entre o pontos i e o ponto j .

Para verificar qual modelo deve ser utilizado na resolução do problema proposto, foram realizados alguns testes. Tais testes visam verificar o tempo gasto em cada modelo e o comportamento dos modelos à medida que o valor de k cresce. Para isso, foram avaliadas as seguintes instâncias: *Ruspint*, *200DATA*, *300p4c1*. As tabelas mostram estes resultados. E assinalam que a variação do k , o valor ótimo para cada k e o tempo de execução. Os modelos foram executados utilizando o software XPRESS. Nos testes de contagem de tempo foram utilizados computadores com processadores Intel Xeon Quad core 3.00 Ghz com 16G de memória RAM.

3. RESULTADOS

No Brasil, há poucos laboratórios de patologia apícola para atender à classe apícola, e ainda é incerto o indicador de perdas de colmeias. Em outros países, um patamar acima de 20% é considerado preocupante. No diagnóstico apícola do estado do Rio de Janeiro, a ocorrência é alta de perdas por doenças de colmeias e há um clima de incerteza quanto ao seu diagnóstico (Figura 1). Observa-se 33,83% dos apicultores (IC 95%, 28,17%; 39,48%) com relato positivo para doença. Esta cifra pode estar subestimada, quando se analisa dentro dos relatos negativos para doença, 49,36% (IC 95%, 41,51%; 57,20%), respostas de incerteza para identificar os sintomas de doenças.

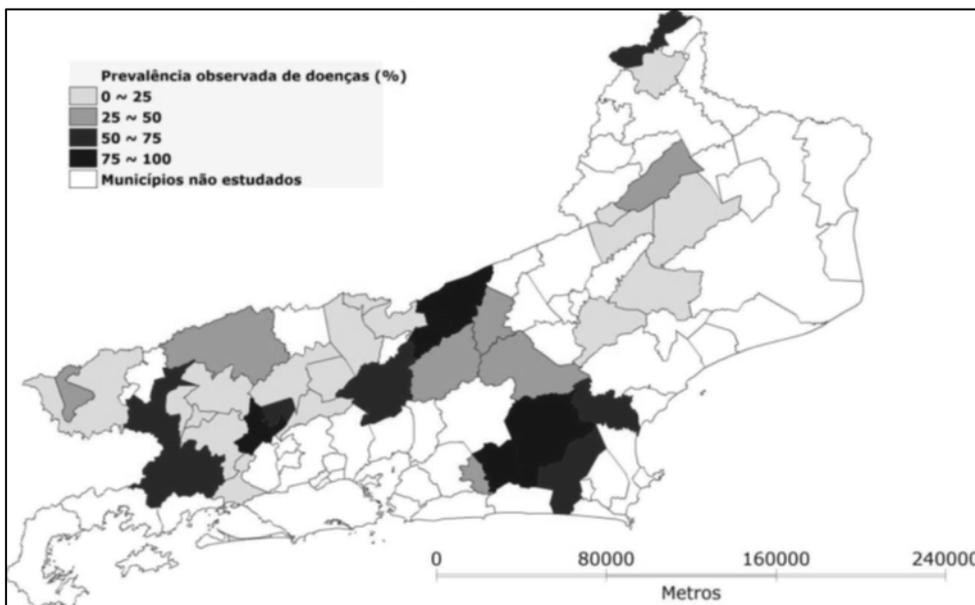


Figura 2: Mapa da prevalência observada de doenças.

Tabela 1: Descrição das instancias

Nome	n ^o pontos	Dimensão
Ruspini	75	R ²
200data	200	R ²
300p4c1	300	R ²

Tabela 2: Testes realizados com o modelo de RAO

Nome	k	ótimo	t(s)
ruspini	K=7	32,984	14163s
	K=6	36,619	1345s
	K=5	40,249	658s
	K=4	47,634	60s
	K=3	88,588	18s
	K=2	102,078	9s
200DATA	K=4	9,257	3688s
	K=3	10,308	344s
	K=2	16,669	243s
300p4c1	K=4	132.966	81356s
	K=3	109,59	1349s
	K=2	365.322	153s

Tabela 3: Testes realizados com o modelo k-medianas

Nome	k	ótimo	t(s)
ruspini	k = 2	2395.8	2s
	k = 3	1619.47	2s
	k = 4	861.47	2s
	k = 5	779.68	2s
	k = 6	714.65	2s
	k = 7	650.84	2s
	k = 8	600.00	2s
	k = 9	555.34	2s
	k =	512.81	2s
	200DATA	k = 2	575.67
k = 3		337.70	4s
k = 4		280.21	4s
k = 5		240.42	3s
k = 6		217.48	3s
k = 7		204.00	3s
k = 8		191.24	3s
k = 9		179.61	3s
k =		168.46	3s

300p4c1	k = 2	16063.8	18s
	k = 3	11001.2	18s
	k = 4	8160.05	17s
	k = 5	7344.33	16s
	k = 6	6563.90	16s
	k = 7	5988.01	16s
	k = 8	5579.9	15s
	k = 9	5286.93	15s
	k =	5019.77	15s
	k =	3412.00	15s
	k =	2706.24	15s
	k =	2274.11	15s
	k =	1963.32	15s

No modelo de RAO, a partir do k acima dos valores definidos na tabela 2, o software demorou mais de um dia e não convergiu para um resultado, e por isso o software foi interrompido. Portanto, os valores de k da tabela, são os únicos valores possíveis executáveis. Em relação ao modelo k-medianas, os valores de k poderiam ser aumentados, porém não seriam relevantes para a comparação. Pelos testes realizados, verificou-se que o modelo k-medianas é mais rápido e consegue executar instâncias de maior cardinalidade, sendo o mais recomendável na condução desta pesquisa.

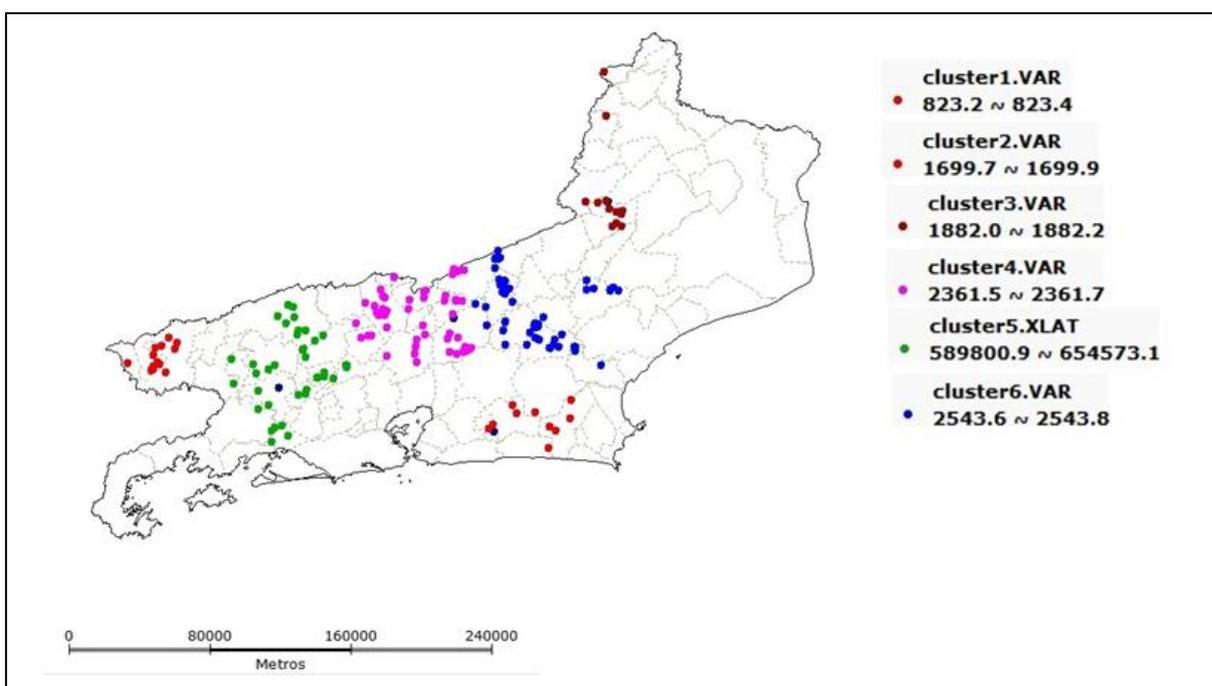


Figura 3: Representação da distribuição espacial dos clusters, indicando os possíveis surtos de doenças apícolas no Estado do Rio de Janeiro.

Observando o mapa do estado do Rio de Janeiro com a representação dos clusters entre as cidades, verifica-se que a localização dos clusters com indicativo de doenças ocorre em todas as mesorregiões do estado do Rio e que há 13 municípios sujeitos à alta prevalência de doenças (acima de 50%): Sapucaia, Mendes, Paracambi, Silva Jardim, Rio Bonito, São José do Vale do Rio Preto, Barra Mansa, Rio Claro, Engenheiro Paulo de Frontin, Casimiro de Abreu, Petrópolis, Araruama e Porciúncula.

Tabela 4: Relação dos municípios e mesorregiões com a distância média dos apiários e prevalências de doenças.

Clusters	Municípios e Mesorregiões	Distâncias médias dos apiários (m)	Prevalência das doenças
1	Itatiaia e Resende (Sul).	823,2 m	0.133
2	Araruama, Rio Bonito, Silva Jardim e Tanguá (Baixada Litorânea).	1699,7 m	0.666
3	Aperibé, Itaocara e Porciúncula (Noroeste e Norte).	1882,0 m	0.086
4	Com. Levy Gasparian, Paraíba do Sul, Paty dos Alferes, Petrópolis, São José do Vale do Rio Preto, Sapucaia, Teresópolis, Três Rios e Vassouras (Centro e Metropolitana).	2361,5 m	0.400
5	Eng ^o Paulo de Frontin, Miguel Pereira, Paracambi, Pinheiral, Piraí, Rio Claro, Barra do Piraí, Barra Mansa, Valença, Vassouras e Volta Redonda, Itaguaí (Sul e Metropolitana).	589800,9 m	0.287
6	Sapucaia, Casimiro de Abreu, Nova Friburgo, São José do Vale do Rio Preto, Sumidouro, Teresópolis e Trajano de Moraes (Sul e Centro).	2543,6 m	0.520

Tabela 5: Relação de produção anual de mel/colmeia e o raio médio dos clusters.

Clusters	Produção anual Mel/Colmeia (kg)	Raio Médio do Cluster (km)
1	13,723 kg	17,61 km
2	10,662 kg	27,01 km
3	13,317 kg	46,79 km

4	7,105 kg	42,50 km
5	9,407 kg	40,50 km
6	7,175 kg	11,59 km

Escala 1: 60000

4. CONCLUSÃO

Em análise geral, verificou-se que os apiários com menor probabilidade de contrair doenças são aqueles onde os materiais apícolas são adquiridos no estado do Rio de Janeiro, existem vizinhos apicultores em um raio de 5 km, não há cultivo agrícola a 5 km do apiário, não é feita divisão artificial de colmeias, a higiene dos apicultores é satisfatória, ocorre florada de alecrim na época de colheita de mel e há assistência técnica para os apicultores.

Medidas para aplicação das boas práticas de produção na apicultura são prementes, especialmente nos municípios de alta prevalência de doenças. Para se desenvolver a apicultura fluminense há necessidade do controle sanitário para prevenir a entrada e a propagação de novas doenças, além do incentivo às pesquisas sobre patologia apícola.

Os resultados obtidos neste estudo podem auxiliar no estabelecimento de estratégias de prevenção e controle fundamentadas na análise dos fatores associados às doenças no estado do Rio de Janeiro e na prevalência de doenças nos municípios estudados.

5. AGRADECIMENTOS

Os autores agradecem ao CNPq e a FAPERJ pelo apoio financeiro concedido e a UFRRJ por toda a ajuda institucional.

6. BIBLIOGRAFIA

[1] CRUZ, Marcelo Dib, e Ochi, Luiz Satoru. "O Problema de Clusterização Automática: Um novo método utilizando ILS". Anais do X Congresso Brasileiro de Inteligência Computacional (X CBIC), Fortaleza-CE, (2011).

[2] CRUZ, Marcelo Dib, and Ochi, Luiz Satoru. "Um Algoritmo Evolutivo com Memória Adaptativa para o Problema de Clusterização Automática". Learning and Nonlinear Models, Volume 8(4), pp. 227-239, (2011).

[3] Freitas D.G. F, Khan A.S., Silva L. M. R. Nível Tecnológico e Rentabilidade de Produção de Mel de Abelha (*Apis mellifera*) no Ceará. *RER*, Rio de Janeiro, vol. 42, nº 01, p. 171-188, (2004).

[4] GESLER, W. The uses of spatial analysis in medical geography: a review. *Social Science and Medicine*, 23:963-973, 1986.

[5] HILLS, M. & ALEXANDER, F. Statistical methods used in assessing the risk of disease near a source of possible environmental pollution: a review. *Journal of the Royal Statistical Society A*, 152:353-363, 1989.

[6] KNOX, E. G. Spatial and temporal studies in epidemiology. In: *Oxford Textbook of Public Health: Methods of Public Health* (W. W. Holland, R. Detels & G. E. Knox, eds.), vol. 2, pp.

95-105, Oxford: Oxford University Press, 1991.

[7] LORENZON, Maria Cristina; TASSINARI, W. S.; KOSHIYAMA; ALMEIDA. Indicadores e Desafios da Apicultura Fluminense-Um Retrato Brasileiro. 1. Ed. Espírito Santo: ABOVE, v. 1000. 272p, (2012).

[8] WERNECK, Guilherme L.; STRUCHINER, Claudio J.. Estudos de agregados de doença no espaço-tempo: conceitos, técnicas e desafios. Cad. Saúde Pública, Rio de Janeiro, v. 13, n. 4, (1997).