

## Algoritmo Genético Aplicado à Análise de Fatores de Risco Associados à Incidência da Malária, no Município de Augusto Corrêa – Pará, Brasil, em 2005

Douglas Gasparetto<sup>1</sup>, Lena Patricia Souza Rodrigues<sup>1</sup>, Diovanni Moraes de  
Araujo<sup>2</sup>, Carlos Roberto Amaral Peixoto<sup>1</sup>, Nelson Fernando de Lisboa Soffiatti<sup>1</sup>,  
Nelson Veiga<sup>1</sup>

<sup>1</sup>Laboratório de Geoprocessamento (LabGeo) – Instituto Evandro Chagas  
(IEC/SVS/MS) Rodovia BR-316 km 7 s/n - Levilândia - 67030-000  
Ananindeua - Pará - Brasil

<sup>2</sup>Universidade Federal do Pará – UFPA  
Rua Augusto Corrêa, 01 - Guamá. CEP 66075-110. Caixa postal 479.  
PABX +55 91 3201-7000 - Belém - Pará - Brasil

{douglaslabgeo, lpsrd13, diovanini, carlosrapeixoto}@gmail.com;  
{nelsonsoffiatti, nelsonveiga}@iec.pa.gov.br

**Abstract.** *This article presents partial results of a model developing project for incidence susceptibility analysis of malaria, in Augusto Correa - Para, Brazil, in 2005. The model was based on the theory of Genetic Algorithms, and had as variables environmental indicator, socioeconomic, epidemiological and entomological. The results enabled us to observe the different variables that influence the incidence of malaria, considering the specifications adopted by the model.*

**Resumo.** *Este artigo apresenta resultados parciais de um projeto de desenvolvimento de um modelo aplicado à análise de fatores de riscos associados à incidência da malária no município de Augusto Corrêa – Pará, Brasil, em 2005. O modelo foi baseado na teoria dos Algoritmos Genéticos, sendo que, teve como variáveis indicadores ambientais, socioeconômicos, epidemiológicos e entomológicos. Os resultados obtidos possibilitaram observar as diferentes variáveis que influenciaram na incidência da Malária, considerando as especificações adotadas pelo modelo.*

### 1. Introdução

Na natureza os seres vivos competem entre si por recursos como água, alimento e território. A partir deste fato, Charles Darwin, em 1859, escreveu em seu famoso livro “A Origem das Espécies” que estas evoluem seguindo um princípio de seleção natural ou de sobrevivência dos mais aptos para enfrentar situações adversas, onde indivíduos desta natureza probabilisticamente têm maior chance de sobreviver.

Baseado nesta teoria, um pesquisador da Universidade de Michigan, John H. Holland propôs, nos anos 60, a criação de algoritmos matemáticos para resolver problemas complexos, denominados então de Algoritmos Genéticos (AGs).

Conforme estudos realizados por Pacheco (1999) e Souza (2004), os AGs são respostas computacionais que buscam a otimização da solução de um determinado problema, a partir do cruzamento de codificações binárias (cromossomos) de diversas variáveis, voltadas para gerar uma possível tomada de decisão. Tais recursos podem ser empregados em diversas áreas do conhecimento, dentre as quais química, biologia, medicina, engenharia, dentre outras.

Outros recursos computacionais para tomada de decisão são as geotecnologias emergentes, que segundo Veiga (2001), Medronho (2004) e Gasparetto (2006) se constituem de sistemas de geoinformação automatizados tendo, coordenadas geográficas como elemento indexador de seus dados, compondo, desta forma, soluções em hardware, software e *peopleware*, que juntos constituem poderosas ferramentas para geração de informações. Conforme Dias (2004), estas ferramentas tem sido apontadas como integradoras de dados ambientais, socioeconômicos e de saúde, permitindo identificar padrões da distribuição de agravos às condições de saúde de populações humanas.

Levando em consideração o exposto anteriormente e partindo do princípio de que a malária é uma doença infecto-contagiosa de etiologia parasitária, com grande expressividade epidemiológica no mundo e que sua incidência está relacionada às características ambientais e relações socioeconômicas presentes no espaço geográfico onde ela ocorre, o presente trabalho teve como objetivo implementar um ambiente computacional baseado em técnicas de AG para analisar os fatores de risco, que determinaram o surto ocorrido no ano de 2005, no município de Augusto Corrêa – Pará, Brasil.

A motivação deste trabalho foi relacionada à possibilidade da compreensão formal, usando um método computacional baseado em uma técnica de Inteligência Artificial (IA), de como a alta incidência desta doença ocorre em função da relação de um conjunto de variáveis que se expressaram de forma diferenciada em um determinado local no espaço geográfico.

## 2 - Metodologia

O método de pesquisa deste trabalho foi estruturado inicialmente com o levantamento do material bibliográfico para construção do referencial teórico sobre Vigilância em Saúde (VS), Malária, Geoprocessamento e Algoritmo Genético e a obtenção das bases de dados dos Sistemas de Informação em Saúde (SIS), das bases cartográficas, epidemiológicas, entomológicas, ambientais e de imagens de satélites, todas de domínio público. Em seguida foi realizada a identificação e o georreferenciamento de dados ecoepidemiológicos coletados em campo, utilizando um receptor do Sistema de Posicionamento Global (GPS), modelo Map 76 CSx, bem como a caracterização ambiental e socioeconômica rápida da área de estudo. Posteriormente, ocorreu a realização do trabalho laboratorial, com o pré-processamento das imagens de satélites utilizando o software ArcGis 9.3, a modelagem do ambiente utilizando o software Visio 2007, conforme fluxo da Figura 1 e a implementação do AG utilizando o software Netbeans 6.5.

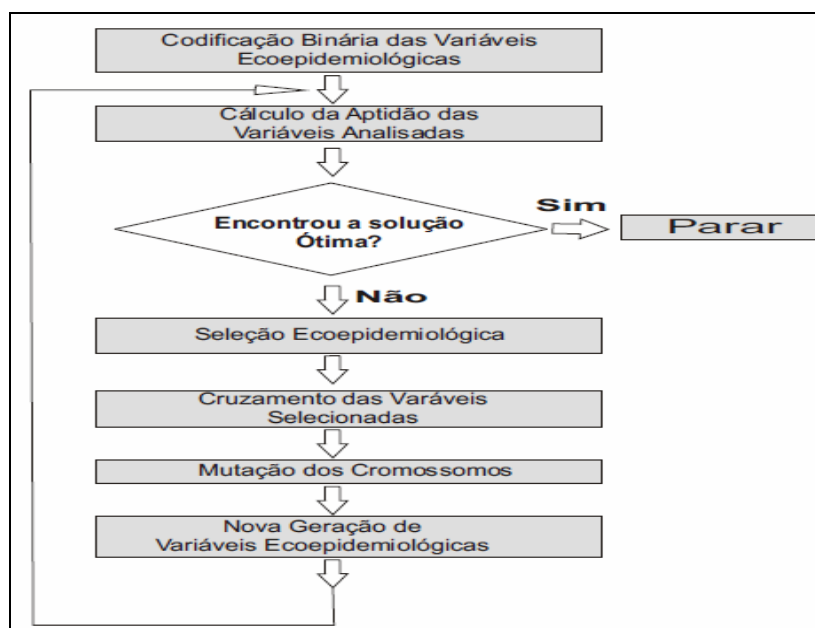


Figura 1: Diagrama do fluxo do AG. Fonte: Adaptado de Poppi et al(1999) Apud Catarina (2005)

Para a construção do AG foi gerado a população inicial de indivíduos (variáveis), composta pelos indicadores ecoepidemiológicos identificados em campo e também pela análise do referencial teórico. Estas variáveis foram codificadas para binário a partir de valores probabilísticos de importância para incidência da malária.

A seleção foi feita conforme as aptidões de cada variável (indivíduo), a maior ou menor probabilidade desta variável influenciar na incidência da malária, conforme preconizado por Lacaz (1997), Forattine (2002), Veiga (2005) e Gasparetto (2006). Atribuindo-se a aptidão de cada indivíduo, foi construída a seleção com um método ponderado, proporcional à probabilidade dos indivíduos.

Após a seleção, foi feito o cruzamento dos pares de variáveis selecionadas e implementado o elitismo da variável de maior risco de infecção por malária. Segundo Filitto (2008) o operador de cruzamento tem a função de combinar os cromossomos dos pais, para gerar os filhos. Contudo, existem vários tipos de operadores de cruzamento, uns desenvolvidos para serem mais genéricos e outros mais adequados a um tipo de codificação de cromossomos. Neste trabalho foi utilizada a técnica de cruzamento com dois pontos de corte, possibilitando, desta forma, aumentar a diversidade e a evolução das variáveis analisadas.

Após o cruzamento foi implementado o operador de mutação, que é responsável pela inserção de pequenas mudanças aleatórias nos cromossomos dos filhos.

Na etapa final foram elaboradas as análises laboratoriais, gerados os relatórios técnicos e posteriormente divulgado os resultados obtidos.

### 3 – Resultados Parciais

Foi gerada a expressão visual dos bancos de dados das informações coletadas em campo a partir da localização espacial dos criadouros, das comunidades na área de estudo e da

caracterização rápida da vegetação, possibilitando o reconhecimento das variáveis que influenciaram na incidência da malária, conforme Figura 2 abaixo.

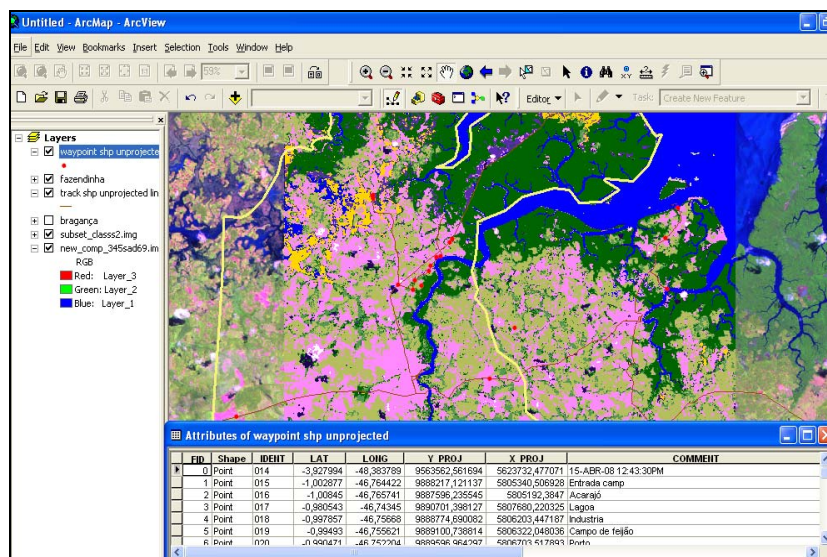


Figura 2. Imagem do interrelacionamento dos bancos de dados de informação. Fonte: LabGeo/IEC/SVS/MS.

Após o desenvolvimento do ambiente de análise ecoepidemiológica baseado em técnicas de AG foi possível observar o comportamento da malária, a partir das possíveis relações que influenciaram no surto da doença, na área de estudo, a partir da utilização dos pesos de cada variável para a incidência da doença (aptidão) e observando sua evolução através do número de gerações (ciclos) que evoluiu, considerando as variáveis analisadas, conforme Figura 3 abaixo.

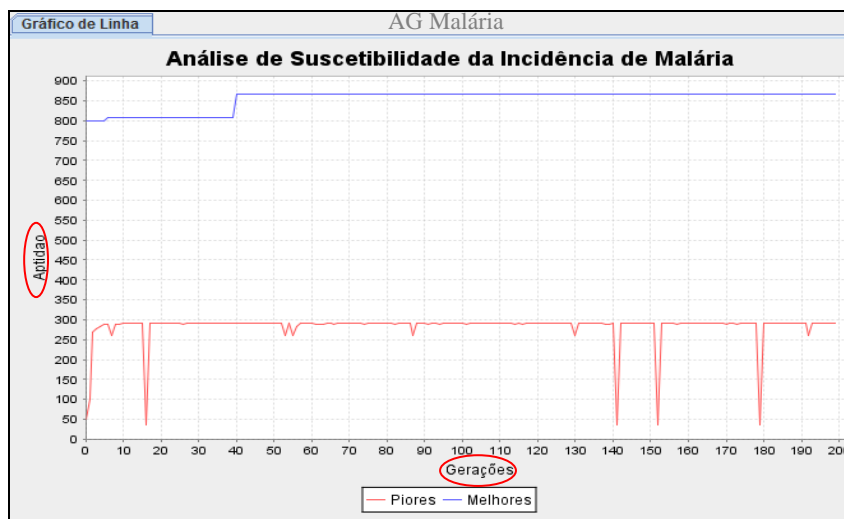


Figura 3 – Análise ecoepidemiológica dos fatores de risco da malária mostrando as aptidões por gerações. Fonte: LabGeo/IEC/SVS/MS.

#### 4 – Considerações Finais

Considerando os resultados obtidos a partir do relacionamento das variáveis identificadas pelas técnicas de geoprocessamento e que foram posteriormente

processadas pelo AG utilizando uma taxa de cruzamento de 90% e de mutação de 1%, foi possível identificar que a partir de 40 ciclos os fatores de risco que influenciaram a incidência da malária convergem com taxa de aptidão 950.

A partir dos testes realizados, com alterações nos valores probabilísticos de importância para a incidência da malária, foi possível observar que a variável vegetação (áreas antropizadas), considerando as características da área de estudo, foi a variável que mais apresentou importância relacionada à incidência da malária. Contudo, mais testes deverão ser feitos para validar este ambiente de análise ecoepidemiológica da incidência deste agravo, baseado em técnicas de AG, utilizando outros métodos de operadores genéticos, aplicados, nesta área de estudo e em outras áreas que apresente fatores de risco de incidência desta doença.

## **5 - Referências Bibliográficas**

- Catarina, A. S. Algoritmos evolutivos aplicados ao processo de análise de dados geográficos. INPE. São José dos Campos - SP, 2005.
- Dias, J. E. et al. Geoprocessamento aplicado à análise ambiental: o caso do município de Volta Redonda – RJ. In.: XAVIER DA SILVA, J.; ZAIDAN, R.T. (org.). Geoprocessamento & análise ambiental: aplicações. p. 368 Rio de Janeiro – RJ, 2004.
- Filitto, D. Algoritmos Genéticos: Uma visão explanatória. Revista Multidisciplinar da UNESP. Saber Acadêmico n. 06. São Paulo - SP, 2008.
- Forattini. Oswaldo Paulo. Culicidologia Médica. Identificação, Biologia, Epidemiologia. vol. 2: Editora da Universidade de São Paulo, São Paulo – SP, 2002.
- Gasparetto, D. SISMAE - Sistema Multifuncional de Avaliação Ecoepidemiológica Aplicado a Prevalência da Malária. Trabalho de conclusão de curso – TCC. CESUPA, Belém - PA, 2006.
- Gomes, J. et al. Computação Gráfica: Imagem. Série Computação e Matemática, SBM/IMPA, Rio de Janeiro – RJ, 2000.
- Lacaz, Carlos da Silva. Temas de medicina: biografia, doenças e problemas sociais. Lemos Editorial. 245p. São Paulo – SP, 1997.
- Medronho, R. A. et al. EPIDEMIOLOGIA. Atheneu ed. São Paulo – SP, 2004.
- Pacheco, M. A. C. et al. Algoritmo Genético: Princípios e Aplicações. ICA: Laboratório de Inteligência Computacional Aplicada. Rio de Janeiro – RJ, 1999.
- Poppi, R. J. et al. Algoritmo Genético em Química. Instituto de Química - Universidade Estadual de Campinas. Campinas – SP, 1999.
- Souza, D. O. de, Algoritmos Genéticos Aplicados ao Planejamento de Transporte Principal de Madeira. Dissertação (Mestrado em Ciências Florestais)-Setor de Ciências Agrárias da universidade Federal do Paraná. 184f. Curitiba - PR, 2004.
- Veiga, N. G. Modelo de Estruturação de Banco Georreferenciado de dados Ambientais, Epidemiológicos e Socioeconômico de Bragança-Pa. Relatório técnico, IEC/SVS/MS. Belém - PA, 2005.