

Modelagem da distribuição do *Anopheles darlingi* em Manaus-AM pelo programa computacional OpenModeller

JAIDSON NANDI BECKER

Instituto Nacional de Pesquisas Espaciais - INPE
Centro de Ciência do Sistema Terrestre - CCST
Pós-Graduação em Ciência do Sistema Terrestre, São José dos Campos, SP
jaidson.becker@inpe.br

RESUMO – O *Anopheles darlingi* é o principal mosquito transmissor da malária no Brasil. Seus criadouros são frequentemente de águas limpas, sombreadas e com baixa correnteza. Portanto, a paisagem pode ser utilizada como indicativo de sua presença. O OpenModeller é um programa computacional para a geração de modelos de distribuição potencial de espécies. Visando a melhor compreensão da dinâmica de transmissão da doença em Manaus, foram gerados 04 modelos de distribuição do vetor.

ABSTRACT - The *Anopheles darlingi* is the main mosquito that transmits malaria in Brazil. His breeding spots often are from waters clean, shaded and slow current. Therefore, the landscape can be used as an indication of their presence. The openModeller is a computer program for generating models of potential distribution of species. Aiming at a better understanding of the dynamics of disease transmission in Manaus, 04 vector distribution models were generated.

1 INTRODUÇÃO

A malária é uma doença infecciosa febril aguda, causada por protozoários do gênero *Plasmodium* e transmitida ao homem, através da picada de mosquitos fêmeas do gênero *Anopheles*. Portanto, sua ocorrência está condicionada à interação de três fatores: o parasito (plasmódio), o hospedeiro (homem) e o vetor (mosquito).

Os mosquitos transmissores da malária compreendem aproximadamente 400 espécies, das quais cerca de 60 ocorrem no Brasil. O *Anopheles darlingi* é o principal vetor (mosquito transmissor) da malária no País. Seus criadouros frequentemente são de águas limpas de baixa correnteza e sombreadas (BRASIL, 2006). Conforme Lambin et al (2010) a paisagem pode ser utilizada como uma indicação (associação) da presença de vetores e hospedeiros da malária.

Neste sentido, este estudo apresenta resultados de modelagem da distribuição espacial do *Anopheles darlingi* para o município de Manaus-AM. Sendo esta informação de forte contribuição para compreensão da dinâmica e dos padrões de disseminação da doença.

2 METODOLOGIA

Este estudo foi aplicado ao município de Manaus/AM por tratar de uma área endêmica para a malária, por apresentar um programa de vigilância epidemiológica atuante e principalmente por possuir uma razoável quantidade de inquéritos entomológicos georreferenciados.

Os dados de presença/ausência do *Anopheles darlingi* foram disponibilizados pelo SIVEP (Sistema de Informação de Vigilância Epidemiológica) através do banco de dados virtual www.saude.gov.br/vetores_malaria (acesso restrito).

A distribuição do vetor foi modelada no programa computacional OpenModeller, sendo utilizados três variáveis ambientais como entrada: Floresta, Vegetação e Hidrologia. As variáveis ambientais foram manipuladas nos programas computacionais SPRING e TerraView de modo a serem compatíveis com os requisitos do programa de modelagem.

As variáveis Floresta e Vegetação foram construídas sobre a partir das classes “floresta”, “floresta secundária”, “pasto limpo”, “pasto sujo”, “pasto em regeneração” e “pasto com solo exposto” do mapa digital TerraClass 2010, disponibilizado pelo INPE. No programa TerraView foi criada uma grade regular com células de 720x720m e atribuído a estas células a distância mínima das variáveis utilizadas (ver figura 1). Ainda no TerraView os valores de distância foram normalizados de 1 a 255 e a grade regular convertida em matriz.

A variável Hidrologia foi construída sobre duas fontes de dados. Do TerraClass retirou-se a classe “hidrologia” (ver figura 1) e de imagens SRTM (Shuttle Radar Topography Mission) foi extraído a rede de drenagem pelo programa

SPRING. As duas feições foram agregadas e o processo metodológico continuou de forma semelhante às variáveis anteriores.

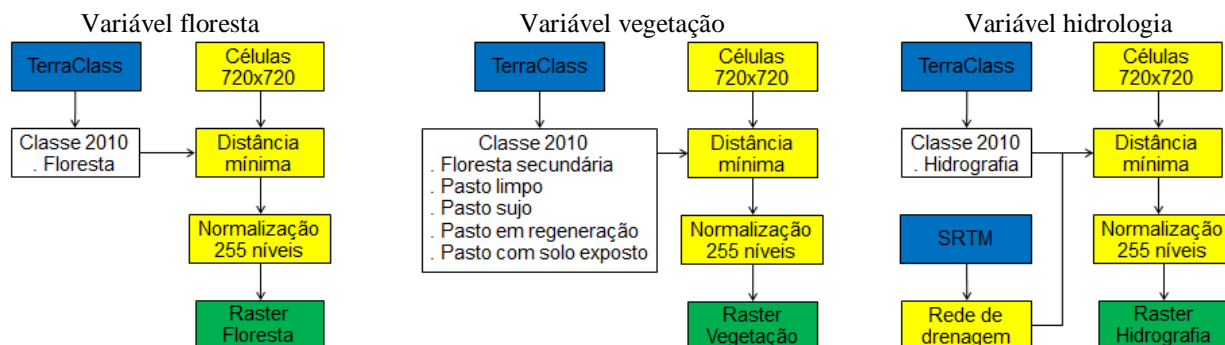


Figura 1 - Diagrama de construção das variáveis.

A modelagem no programa OpenModeller obedeceu os passos apresentados no diagrama da figura 2. Foram utilizados dois algoritmos para a modelagem: por distância euclidiana e por distância de Mahalanobis. Dois parâmetros de vizinhança diferentes também foram usados: por vizinhança única e por vizinhança global. O que gerou quatro resultados diferentes de distribuição espacial para o *Anopheles darlingi*.

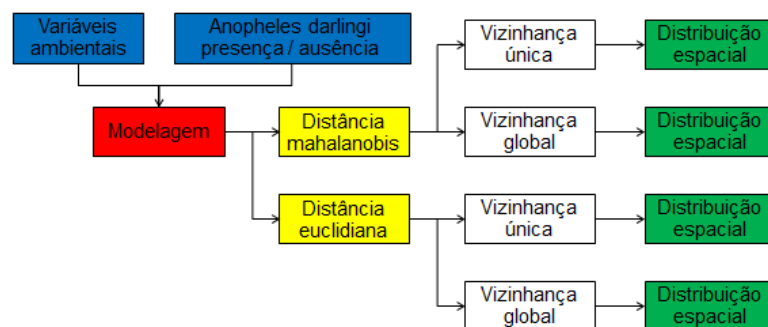


Figura 2 - Diagrama de modelagem.

Os dados de presença/ausência do *Anopheles darlingi* podem ser visualizados no mapa da figura 3, onde os pontos verdes representam localidades com registro de captura negativa (ausência do mosquito) e os pontos vermelhos são localidades com captura positiva (presença do mosquito). Para o período de 01 de janeiro á 30 de junho de 2011 foram registradas 35 localidades com presença do *An. darlingi* e 18 localidades com ausência.

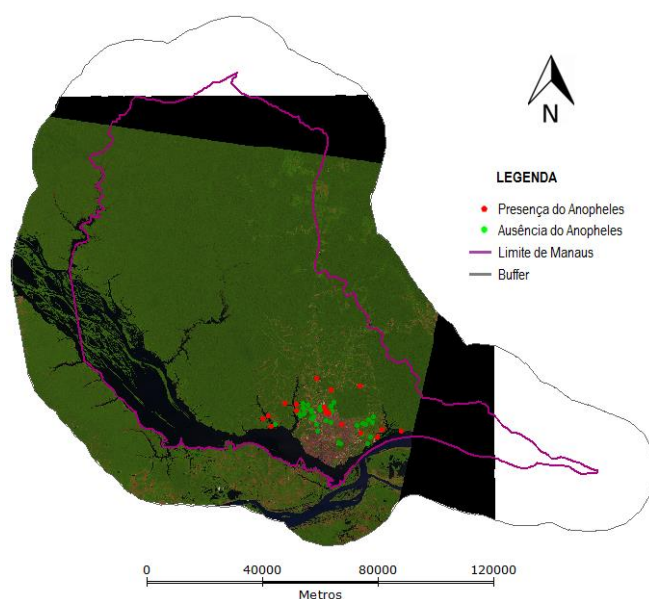


Figura 3 - Mapa do registro de captura do *Anopheles darlingi* para Manaus/2011.

As variáveis Floresta, Vegetação e Hidrografia podem ser visualizadas nos mapas da figura 4, 5 e 6 respectivamente. As variáveis foram construídas conforme descrito na metodologia e a sua representação em raster informa a distância mínima em níveis de cinza (1 a 255). Desta forma a cor escura (255) representa o local (a célula), existente no espaço coberto pelo mapa, de maior distância da paisagem, ou classe de cobertura da terra, referente. Enquanto a cor clara (01) representa a célula com esta paisagem presente (distancia zero). As demais células (entre 01 e 255), coerentemente, representam as distâncias intermediárias.

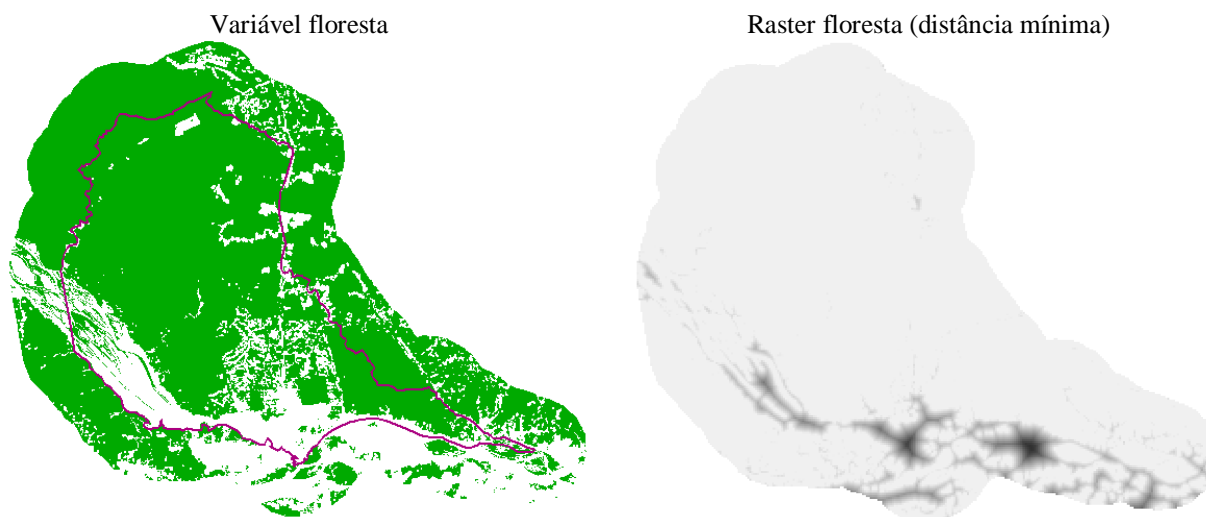


Figura 4 - Variável floresta.

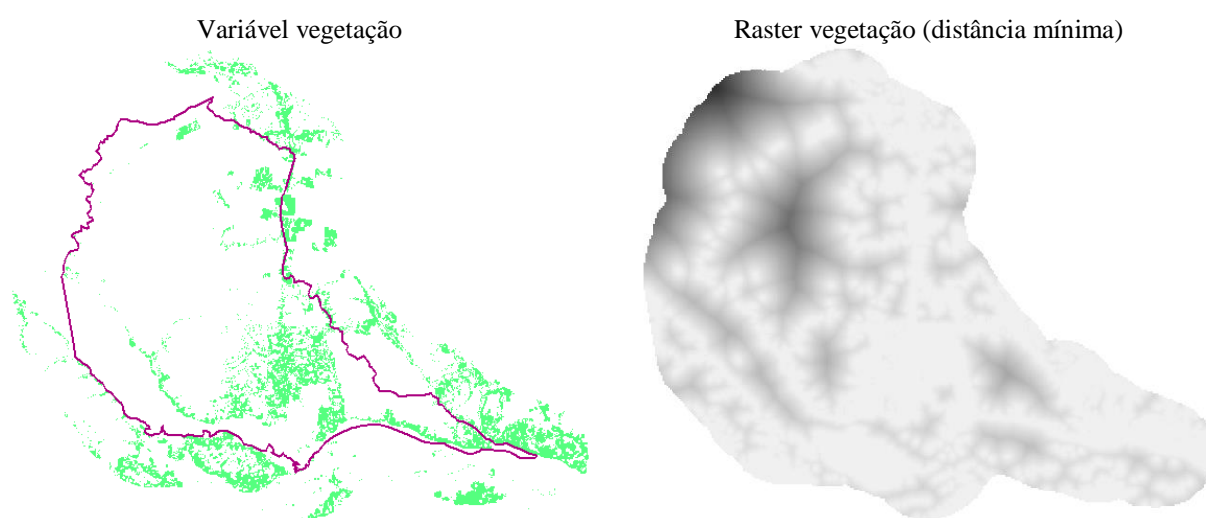


Figura 5 – Variável vegetação.



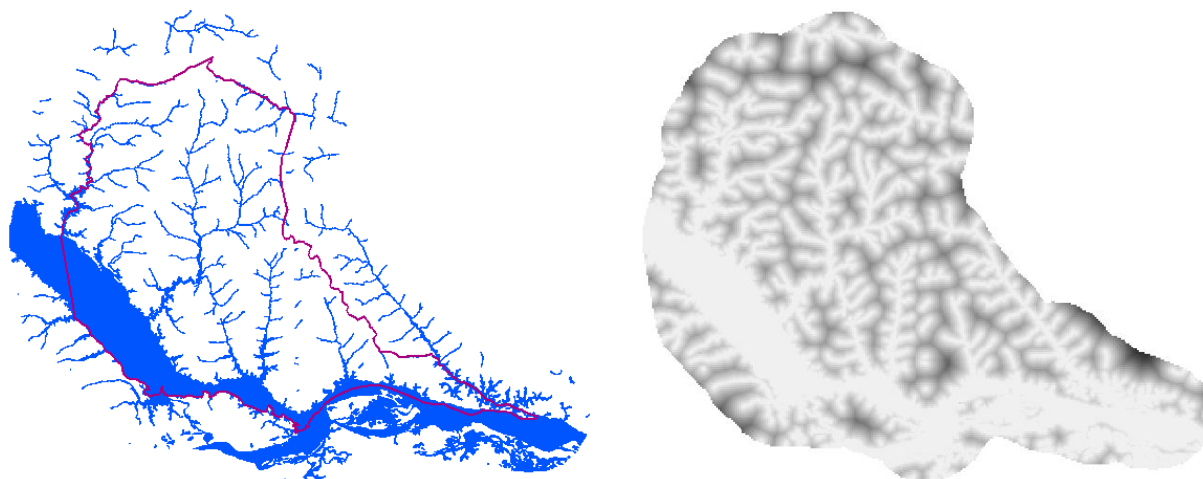


Figura 6 - Variável hidrografia.

3 RESULTADOS

Na sequência (figuras 07 a 10) estão apresentados os quatro modelos de distribuição espacial do *Anopheles Darling* para o municípios de Manaus gerados no programa OpenModeller. Nos mapas das referidas figuras, a cor azul (fria) indica baixa probabilidade de presença do *An.darlingi*, enquanto a cor vermelha (quente) indica alta probabilidade.

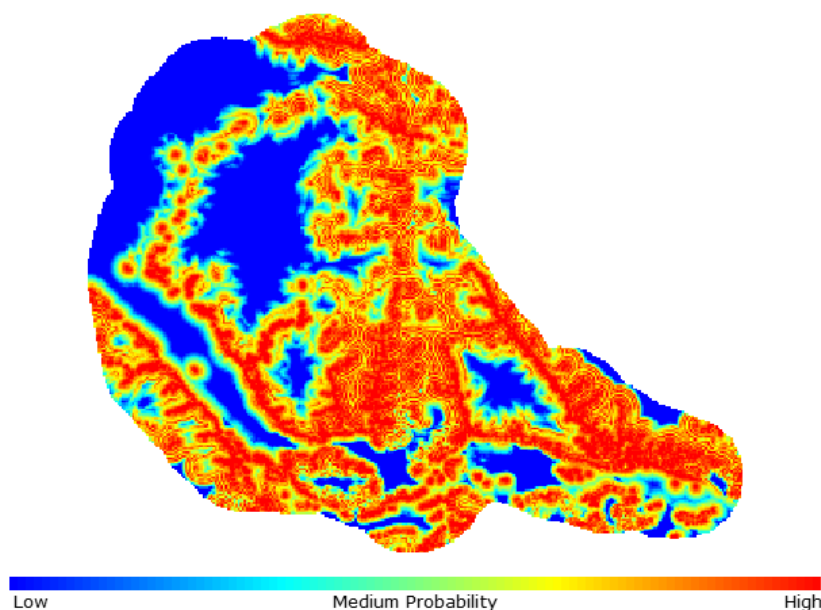


Figura 7 - Modelo de distribuição utilizando o algoritmo de distância euclidiana e o parâmetro de vizinhança única.

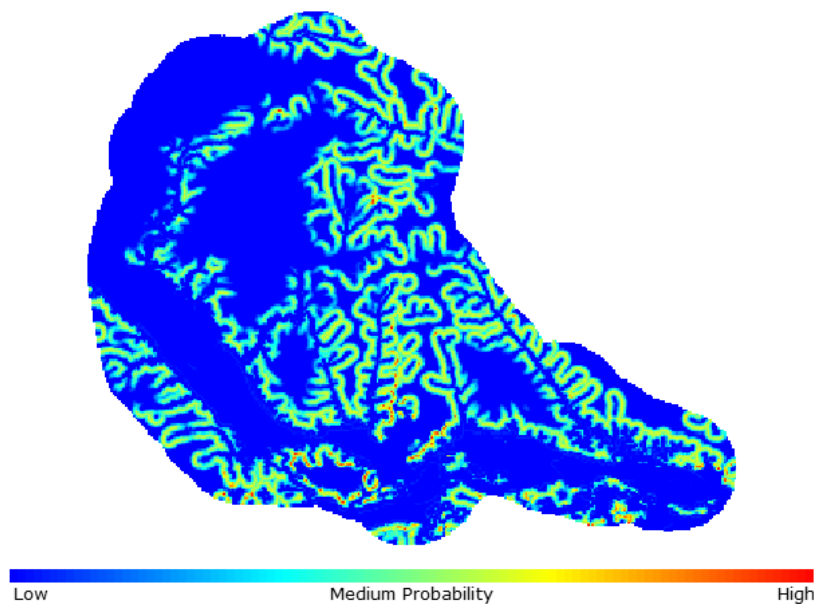


Figura 8 - Modelo de distribuição utilizando o algoritmo de distância euclidiana e o parâmetro de vizinhança global.

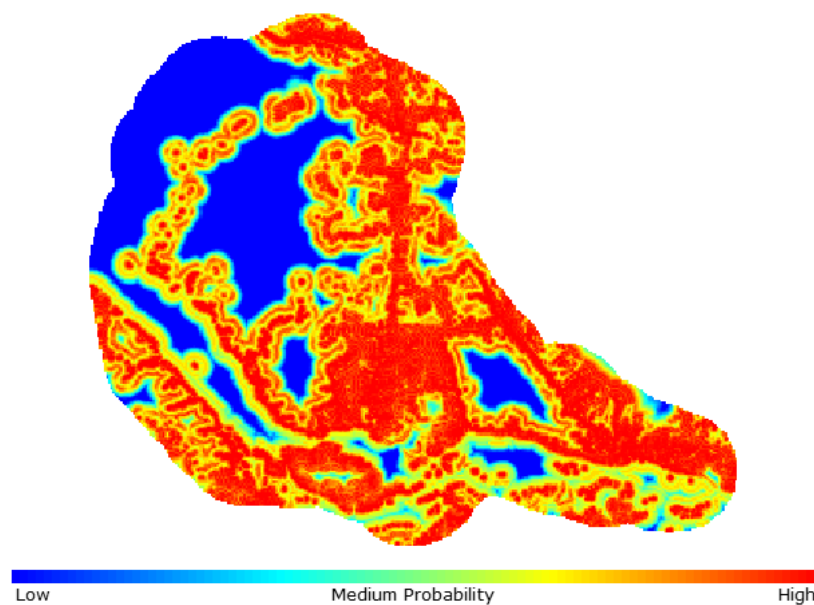


Figura 9 - Modelo de distribuição utilizando o algoritmo de distância Mahalanobis e o parâmetro de vizinhança única.

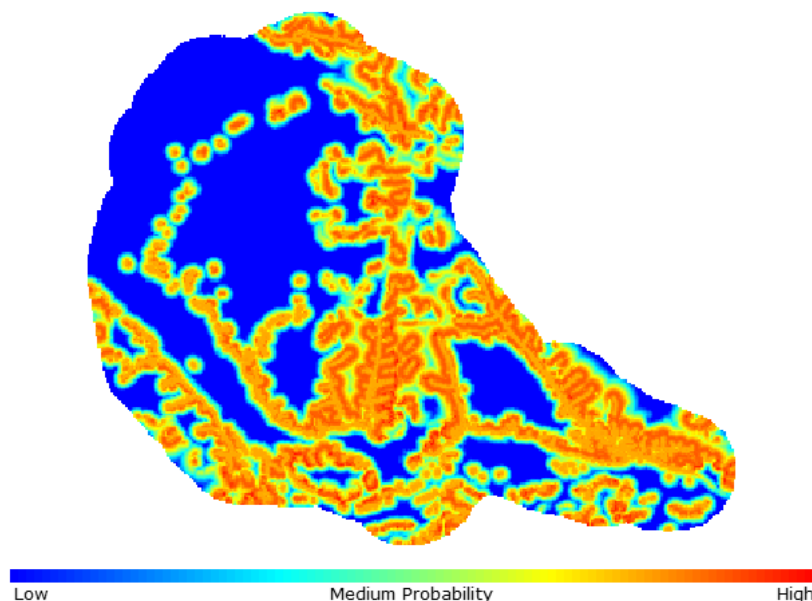


Figura 10 - Modelo de distribuição utilizando o algoritmo de distância Mahalanobis e o parâmetro de vizinhança global.

Os gráficos da Curva ROC (*Receiver Operating Characteristic*) para os quatro modelos de distribuição do *An. darlingi* estão apresentados abaixo (figura 11 e 12). A curva ROC é um teste de validação que quantifica o poder discriminativo de um procedimento para determinada análises, no caso a predição de distribuição do *An. darlingi*. De um modo geral, valores de AUC (Area Under Curve) acima de 0,5 são considerados bons, sendo melhores quanto mais próximos de 1,0.

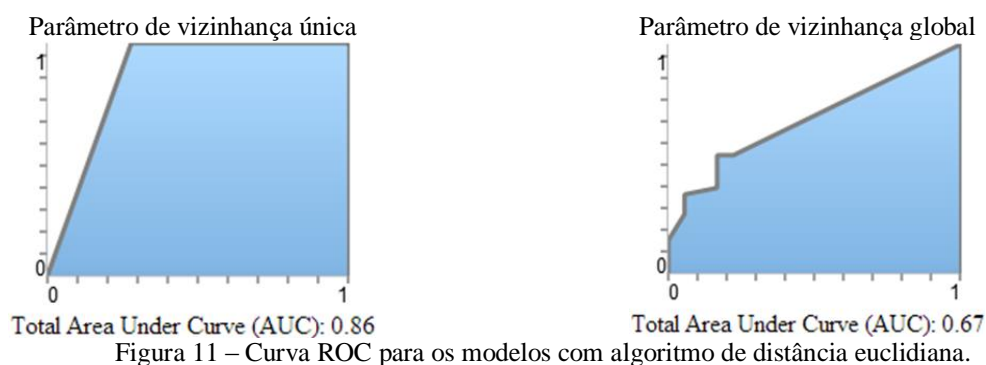


Figura 11 – Curva ROC para os modelos com algoritmo de distância euclidiana.

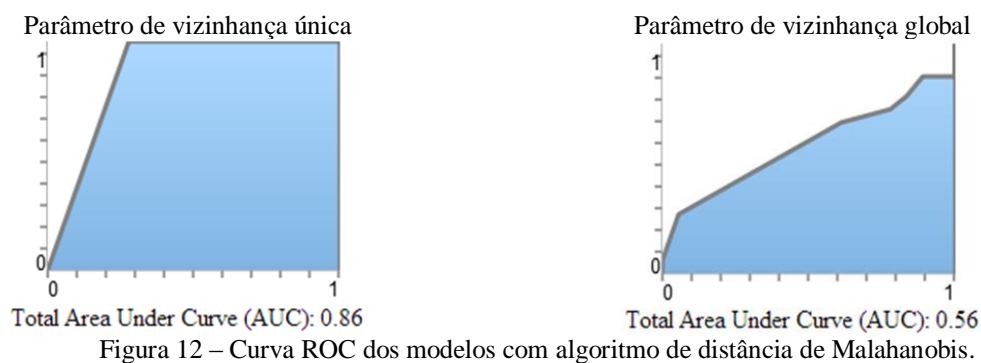


Figura 12 – Curva ROC dos modelos com algoritmo de distância de Malahanobis.

4 DISCUSSÕES

O programa OpenModeller não trabalha com dados de abundância de espécies, portanto o resultado final é a probabilidade da presença da espécie no território e não a quantidade (abundância) de indivíduos, o que seria interessante ao estudo.

Além dos resultados apresentados, o modelo também foi testado sem os dados de ausência. Observou-se que o resultado final sem estes dados não apresentam modificações. A diferença está em que com os dados de ausência há um aumento do valor da curva ROC e o programa gera uma matriz de confusão e consequente Índice Kappa na tentativa de melhor validar o modelo.

Demais métricas para o algoritmo de distância (Gower e Shebyshev) também foram testadas, assim como o algoritmo GARP (Genetic Algorithm for Rule-set Production). Estas métricas apresentaram resultados semelhantes à distância euclidiana. Pelo programa tratar-se de uma caixa-preta, onde não se conhece e nem se tem acesso às formulações matemáticas, optou-se trabalhar somente com a distância euclidiana e a de Mahalanobis por serem mais difundidas e de fácil compreensão.

Nos modelos gerados utilizou-se o parâmetro de distância máxima igual a 10%. O valor de distância máxima define o desvio considerado aceitável sobre os valores das variáveis ambientais utilizadas. Como as variáveis ambientais foram normalizadas de 1 a 255, a variação aceita é de 25,5.

Os dados de presença e ausência do vetor utilizados foram referentes a um período de 6 meses. Na modelagem optou-se por não considerar-se a escala temporal, embora a mesma tenha sido testada. Esta decisão foi tomada pela captura dos vetores não terem seguido um padrão espacial pré-determinado (mesmas localidades de captura em todos os meses). O que neste caso, prejudicaria o resultado.

5 CONCLUSÕES

Os modelos com vizinhança única apresentaram melhor curva ROC (0,86 tanto para o algoritmo de distância euclidiana quanto para a distância de Mahalanobis). Na vizinhança única a modelagem é realizada em relação aos valores das variáveis ambientais referentes ao ponto de presença mais próximo. Portanto o resultado final é algo mais adaptado localmente. Isto gera um resultado de aparência mais segmentada, gerando bordas entre os níveis de probabilidade.

Os modelos com vizinhança global apresentaram menor valor de curva ROC (0,67 para distância euclidiana e 0,56 para a de Mahalanobis). A vizinhança global utiliza o valor médio das variáveis ambientais de todos os pontos de presença do Anopheles. O resultado é a distribuição da probabilidade de forma mais suavizada e homogênea.

Para a vizinhança única obteve-se uma maior área de alta probabilidade do que para vizinhança global. Portanto, a utilização da média global tornou o algoritmo mais restritivo. Isto indica que há um significativo desvio padrão entre os valores das variáveis ambientais associadas aos pontos de presença do Anopheles. Nota-se que para o algoritmo de Mahalanobis a restrição foi menor do que para a distância euclidiana.

Nos os quatro modelos gerados observa-se uma alta probabilidade entorno da variável hidrografia e dentro da variável vegetação. No modelo, a variável floresta acaba atuando como um restritivo da presença do mosquito, o que não acontece no mundo real. Este resultado se deve ao fato dos dados estarem possivelmente viciados. As capturas de alados ocorreram todas em local peridomiciliar (ao céu aberto e próximo de domicílios). Nenhuma captura ocorreu em área tida como floresta prejudicando o resultado final.

Deve-se atentar que a captura de alados presta-se a fim de vigilância epidemiológica não estando no planejamento da mesma a modelagem da distribuição das espécies. Futuramente é possível que a mesma seja planejada, incluindo locais de captura com cobertura e uso da terra heterogêneos.

Por fim, dos resultados apresentados, o que melhor se assemelha a realidade é o modelo de Mahalanobis com vizinhança global, mesmo este possuindo menor curva ROC. Acredito que as variáveis ambientais obtiveram melhores resultados pela análise de sua covariância (considerando a qualidade dos dados de amostragem). E por tratar-se a área de estudo de uma região pequena, a vizinhança global é a mais indicada. Deve-se lembrar que tratando-se de modelagem o conhecimento de quem a faz é fundamental para seu resultado final, e aqui coloco as minhas conclusões pessoais baseadas em meu conhecimento sobre a dinâmica de transmissão da malária.

REFERÊNCIAS

BRASIL. Ministério da Saúde. **Ações de controle da malária**. Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Vigilância Epidemiológica. Brasília/DF, 2006.

BRASIL. Ministério da Saúde. **Guia de vigilância epidemiológica**. Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Vigilância Epidemiológica. Brasília/DF, 2010.

LAMBIN, E. F. TRAN, A. VANWAMBEKE, S. O. LINARD, C. SOTI, V. **Pathogenic landscapes: Interactions between land, people, disease vectors, and their animal hosts**. International Journal of Health Geographics, 2010.