



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES  
**INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS**

**USO DE GAM PARA ANÁLISE DE POTENCIAL  
PRESENÇA DE CARAMUJOS *BIOMPHALARIA* NA  
RGI DE OURINHOS - SP**

Vivian Alessandra Ferreira da Silva

Trabalho final da disciplina de  
Análise Espacial de Dados Geográficos do Curso de Mestrado em Sen-  
soriamento Remoto



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES  
**INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS**

# USO DE GAM PARA ANÁLISE DE POTENCIAL PRESENÇA DE CARAMUJOS *BIOMPHALARIA* NA RGI DE OURINHOS - SP

Vivian Alessandra Ferreira da Silva

Trabalho final da disciplina de  
Análise Espacial de Dados Geográficos do Curso de Mestrado em Sen-  
soriamento Remoto

2021

## LISTA DE FIGURAS

	<u>Pág.</u>
2.1 Distribuição espacial dos hospedeiros intermediários da Esquistossomose	5
3.1 Mapa de localização da Região Geográfica Imediata (RGI) de Ourinhos	11
3.2 Distribuição espacial dos dados considerados na análise. . . . .	13
4.1 Resultado do modelo. . . . .	15
4.2 Gráfico da curva ROC calculada para o modelo . . . . .	16
4.3 Mapa de potencial presença de caramujos <i>Biomphalaria</i> . . . . .	16

## LISTA DE TABELAS

	<u>Pág.</u>
3.1 Variáveis incluídas no GAM. . . . .	13

## SUMÁRIO

	<u>Pág.</u>
<b>1 INTRODUÇÃO</b> . . . . .	<b>1</b>
<b>2 REFERENCIAL TEÓRICO</b> . . . . .	<b>3</b>
2.1 Esquistossomose: expansão, transmissão . . . . .	3
2.2 <i>Biomphalaria</i> : caramujo hospedeiro intermediário . . . . .	4
2.3 Esquistossomose e disponibilidade hídrica . . . . .	5
2.4 Controle e vigilância da Esquistossomose . . . . .	6
2.5 Epidemiologia espacial . . . . .	7
2.6 GAM . . . . .	8
2.7 Curva ROC . . . . .	9
<b>3 MATERIAIS E MÉTODOS</b> . . . . .	<b>10</b>
3.1 Área de estudo . . . . .	10
3.2 Metodologia . . . . .	10
3.3 Definição, criação e organização das variáveis . . . . .	11
3.4 Ajuste do Modelo . . . . .	14
<b>4 ANÁLISE E RESULTADOS</b> . . . . .	<b>15</b>
<b>5 CONCLUSÕES</b> . . . . .	<b>18</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b> . . . . .	<b>19</b>

# 1 INTRODUÇÃO

Os caramujos do gênero *Biomphalaria* são aquáticos e habitam coleções de água doce com pouca correnteza ou parada. Possuem concha discoidal em espiral plana, com os lados aproximadamente paralelos. Geralmente os moluscos sobrevivem na natureza durante um ano e sua persistência nos focos relaciona-se com seu ritmo de reprodução, que depende de diversos fatores ecológicos que podem influenciar na fecundidade e viabilidade dos ovos. Os caramujos têm importância médica por serem hospedeiros intermediários do *Schistosoma mansoni*, parasita causador da esquistossomose em humanos no Brasil. Existem dez espécies e subespécies de *Biomphalaria* no Brasil, dentre as quais três são hospedeiros intermediários: *Biomphalaria glabrata*, *Biomphalaria straminea* e *Biomphalaria tenagophila* (CARVALHO, 2014; SAÚDE, 2009).

A doença, popularmente conhecida como barriga d'água, é uma doença infecciosa, negligenciada, parasitária, tropical (FIOCRUZ, 2014). Dezenove estados brasileiros são considerados áreas endêmicas da esquistossomose, ou seja, são locais onde o ciclo de transmissão é estabelecido (KATZ, 2018). Sua transmissão se dá em um ciclo complexo relacionado à disponibilidade de água no ambiente. As doenças parasitárias e as doenças tropicais negligenciadas, categorias em que a esquistossomose se encaixa, estão entre os agravos em saúde que apresentam comportamento epidemiológico mais relacionado à existência de elementos do ambiente que influenciam, em graus variados, a sua ocorrência e a dinâmica de vida dos parasitas, vetores e hospedeiros (CONFALONIERI, 2005; HAMM et al., 2015; CASTAGNA, 2010).

Este trabalho teve como objetivo empregar o Modelo Aditivo Generalizado (GAM) para espacializar o potencial de presença de caramujos do gênero *Biomphalaria* na Região Geográfica Imediata de Ourinhos, no estado de São Paulo, utilizando variáveis relacionadas à disponibilidade de água no ambiente e dados de presença e não-presença dos caramujos na área de estudo.

Os capítulos restantes deste trabalho estão organizados da seguinte maneira:

- Capítulo 2: Este capítulo trata do referencial teórico que baseia o trabalho. Ele se inicia com a expansão e transmissão da Esquistossomose no Brasil, e descrição do hospedeiro intermediário. Em seguida, faz-se uma breve revisão sobre a relação entre a Esquistossomose e disponibilidade hídrica, e sobre controle e vigilância da doença. Posteriormente a Epidemiologia Espacial é apresentada de forma resumida, assim como o do Modelo Aditivo

Generalizado (GAM) e a curva ROC, métodos empregados no capítulo seguinte.

- Capítulo 3: Neste capítulo é realizada uma breve descrição da área de estudo e da metodologia do trabalho, que inclui explicação sobre as variáveis e ajuste do modelo.
- Capítulo 4: Neste capítulo a análise e os resultados obtidos são discutidos.
- Capítulo 5: Com base nas análise realizada no Capítulo 4, neste capítulo serão apresentadas as conclusões obtidas, como também algumas sugestões para trabalhos futuros.

## 2 REFERENCIAL TEÓRICO

### 2.1 Esquistossomose: expansão, transmissão

Entende-se que a esquistossomose foi introduzida no Brasil através de população africana escravizada na produção canavieira no nordeste brasileiro. Lá havia condições bioecológicas para que se completasse o ciclo evolutivo do parasita. Movimentos migratórios nos séculos XVIII, XIX e XX, relacionados ao desenvolvimento das atividades econômicas do país, provavelmente tornaram possível a reunião de indivíduos infectados e moluscos hospedeiros intermediários do parasita em várias partes do Brasil. De forma geral, a esquistossomose em território brasileiro acompanhou as correntes de migração interna, sendo condicionada pela presença do molusco hospedeiro intermediário. Na década de 1950, após inquérito de nível nacional, delimitou-se as zonas endêmicas, índices de prevalência e a gravidade da expansão da esquistossomose no Brasil. Nas duas décadas seguintes, ao lado do aumento na quantidade de infectados, percebeu-se a tendência à expansão territorial das áreas endêmicas (CHIEFFI; WALDMAN, 1988; CASTAGNA, 2010). Estimava-se existirem cerca de 6,3 milhões de portadores do parasita em 18 estados brasileiros em 1997 (KATZ; PEIXOTO, 2000).

A transmissão da doença pode ser entendida num ciclo onde todos os estágios dependem da disponibilidade de água doce. Os ovos do *schistosoma mansoni* presente nas fezes humanas infectadas eclodem quando entram em contato com água doce à temperatura ideal de 28°C. As larvas ciliadas (miracídios) liberadas com a eclosão têm tempo de vida de cerca de 8 horas e deslocam-se no meio aquático em busca de um hospedeiro intermediário susceptível (caramujo *Biomphalaria*), que aloja o parasita em sua fase larvária ou de reprodução assexuada. Em contato com o hospedeiro, as larvas penetram no molusco em um processo que dura cerca de 15 minutos. Após um período aproximado de 25 a 30 dias, quando as larvas atingem o ponto de cercárias, o molusco inicia a liberação de grande quantidade de larvas infectantes na água. Nesse estágio o parasita é capaz de penetrar no corpo humano dos indivíduos que entram em contato com água contaminada e continuar seu desenvolvimento até atingir a maturidade cerca de 28 dias depois. Hospedado-se no intestino do ser humano, o verme libera, em média, 300 ovos por dia, o que reinicia o ciclo de transmissão da doença (BRASIL, 2008; PESSOA; MARTINS, 1988).

A sintomatologia clínica da esquistossomose depende de seu estágio de evolução no hospedeiro definitivo, o ser humano. A fase aguda pode ser assintomática, sendo frequentemente acidentalmente revelada em exames médicos de rotina, devido a al-

terações nos exames laboratoriais (ANDRADE, 2004). A forma aguda sintomática se apresenta em quadro clínico único de sintomatologia parca, com sintomas dermatológicos e prurido, podendo evoluir para a forma aguda febril. Após seis meses de infecção há o risco do quadro clínico evoluir para a fase de esquistossomose crônica, quando pode haver a forma intestinal (assintomática ou acompanhada de sintomas como diarreia, hepatomegalia, hepatoesplenomegalia e epigastralgia) ou até a forma hepatoesplênica descompensada grave, que apresenta fígado volumoso ou contraído devido à fibrose, ascite, anemia, desnutrição e hiperesplenismo (CASTAGNA, 2010).

## 2.2 *Biomphalaria*: caramujo hospedeiro intermediário

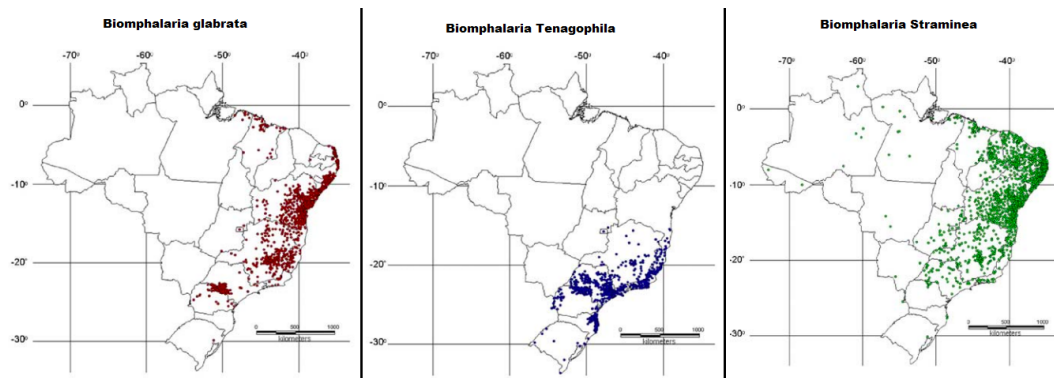
Os caramujos do gênero *Biomphalaria* podem viver em coleções de água natural ou artificial de água doce e lântica, como lagoas, lagos, poças, cisternas, pântanos, remansos de rios, riachos, canais de irrigação e drenagem, barragens, plantações ou em qualquer área inundada (PIERI, 1995). Normalmente os habitats apresentam rica microflora e matéria orgânica, pouca turbidez, luminosidade, pH entre 6 e 8, teor de sal abaixo de 0,3% e temperatura em torno de 20 a 25° C. A vegetação associada aos criadouros é essencial para a manutenção de um microclima que é protegido da radiação solar, altas temperaturas e correntes de água (BRASIL, 2008).

A vida média dos caramujos *Biomphalaria* é de aproximadamente um ano e sua rápida reprodução, possibilitada por serem hermafroditas e conseguirem realizar autofecundação, é um dos fatores de sobrevivência e permanência das espécies do gênero. Os caramujos têm grande capacidade de adaptação às condições adversas, o que é um aspecto ecológico importante para a manutenção da população de moluscos do gênero em condições desfavoráveis. Eles podem sobreviver fora da água por longos períodos ao retraírem suas partes moles dentro da concha impedindo a dessecação por formação de uma película impermeável que evita a perda de água. Também são capazes de se enterrar em uma fina camada de solo em períodos de seca e entrar em hibernação em situações de falta de alimento, diminuição ou aumento da temperatura (BRASIL, 2008; CARVALHO, 2014).

Os hospedeiros intermediários *Biomphalaria glabrata*, *Biomphalaria straminea* e *Biomphalaria tenagophila* têm distribuições espaciais diferentes, mas a ocorrência do *B. glabrata* está quase sempre associada à distribuição da esquistossomose (BRASIL, 2008). A Figura 2.1 mostra a distribuição das três espécies no Brasil.



Figura 2.1 - Distribuição espacial dos hospedeiros intermediários da Esquistossomose



Fonte: Adaptado de CARVALHO O. S.; CALDEIRA (2004).

### 2.3 Esquistossomose e disponibilidade hídrica

De acordo com Teles (2005), a distribuição da transmissão da esquistossomose mansoni no estado de São Paulo e no restante do país não é homogênea e não parece seguir um padrão de distribuição espacial. Entretanto, nos locais onde ela ocorre é evidente a presença na paisagem de um ou mais fatores ecológicos favoráveis à manutenção do ciclo de vida completo do parasita. Para avaliar as possíveis relações entre os elementos da paisagem e a vulnerabilidade à transmissão de esquistossomose mansoni, Castagna (2010) identificou 61 características do meio ambiente que podem favorecer a instalação e reprodução de *Biomphalaria* e a transmissão do *Schistosoma mansoni*. Os fatores considerados estão relacionados a agricultura anual e de subsistência, agricultura semianual ou perene, áreas permanentes ou eventualmente alagadas, águas e valas para irrigação de cultivos, meandros retificados, ambientes próximos a cursos d'água propícios à poluição por dejetos humanos, lagoas de piscicultura, lagoa assoreada ou pouco profunda, lagoas ou curso d'água utilizados para esporte, pesca, lavagens de utensílios e banho, várzea, presença de nascentes ou bicas, depósitos arenosos nas bordas dos rios desprovidos de vegetação, grandes depósitos de entulho e lixo, atividades humanas nas bordas de curso de água, área de preparo agrícola, cobertura vegetal.

De maneira similar, Anaruma et al. (2007) identificaram como fatores indutores para esquistossomose algumas características dos corpos hídricos e certas formas de uso e ocupação das margens de rios e lagoas além dos descritos por Castagna (2010), tais como: meandros e canais naturais; lagoas com presença de vegetação aquática

flutuante ou macrófitas fixas no fundo (taboa); terras baixas alagadiças, vegetação ciliar e fragmentos florestais. Para os autores, cada indutor participa positiva ou negativamente na evolução da transmissão da esquistossomose por meio de redes de interação.

[MOURA e al. \(2005\)](#) observaram a importância do mapeamento dos recursos hídricos para a caracterização de áreas de risco, uma vez que esta variável é fundamental para o conhecimento da distribuição do hospedeiro intermediário. Os autores procuraram associar o risco para esquistossomose às áreas de maior concentração de recursos hídricos a partir de mapeamento da hidrografia através de um modelo digital de elevação de terreno obtido a partir de imagens de radar. Essa associação da transmissão da doença com a concentração de recursos hídricos favorece a delimitação da área de estudo respeitando-se os limites das bacias hidrográficas, inclusive porque a vulnerabilidade dessas áreas frente as atividades humanas refletem com precisão desequilíbrios relacionados a mudanças no uso da terra ([CASTAGNA, 2010](#)).

## **2.4 Controle e vigilância da Esquistossomose**

Até a década de 70, o combate à esquistossomose no Brasil tinha como objetivo principal o controle da transmissão, sendo a principal medida a redução das populações dos moluscos hospedeiros intermediários. Na década seguinte, com o desenvolvimento e uso de drogas quimioterápicas mais eficazes, baratas e seguras, o objetivo passou a ser o controle da morbidade da doença. Posteriormente, medidas preventivas como a educação em saúde e o saneamento tiveram papel importante na redução dos números da enfermidade ([BRASIL, 2008](#)).

Os métodos de controle dos hospedeiros intermediários podem ser classificados em biológico, físico e químico. O biológico envolve a introdução de organismos com ação predadora, competidora, parasitária ou patogênica sobre os moluscos hospedeiros. Outros moluscos, peixes, e outros animais competidores como camarões, insetos e sanguessugas podem ser utilizados como competidores ou controladores das populações de caramujos. No entanto, desde a década de 90, a introdução de espécies exóticas para controle dos moluscos tem sido desencorajada pela Organização Mundial da Saúde. O controle físico, capaz de impedir o desenvolvimento e manutenção das populações de moluscos, consiste na manipulação de fatores do meio, sendo o saneamento hídrico e o manejo ambiental para o controle de vetores exemplos da prática. Objetiva-se a eliminação de criadouros através de aterramento de coleções hídricas, por exemplo, quando a medida é tecnicamente recomendada. A drenagem, retificação de leitos, o revestimento e a canalização dos cursos d'água também podem ser

úteis, bem como a limpeza e a remoção da vegetação aquática. O controle químico consiste na aplicação de produtos tóxicos (moluscidas) nos criadouros (BRASIL, 2008).

Outras ações dirigidas aos hospedeiros intermediários são indicadas em situações de levantamento de áreas ainda não trabalhadas, investigação de focos, inclusive com trabalhos em campo, e em áreas com registro de altas prevalências. Estão entre elas a pesquisa de coleções hídricas para determinação do seu potencial de transmissão; e medidas de saneamento ambiental para dificultar a proliferação e o desenvolvimento dos hospedeiros intermediários, bem como impedir que o homem infectado contamine as coleções de águas com ovos de *Schistosoma mansoni* (SAÚDE, 2009).

Como a maior parte dos casos é assintomática e crônica, os programas de educação em saúde fazem parte das estratégias de controle da esquistossomose. Utilizam-se técnicas pedagógicas veiculadas em meios de comunicação de massa para difundir o conhecimento sobre a doença e estimular as pessoas a adotarem atitudes e práticas preventivas que modificam as condições favorecedoras e mantenedoras da transmissão. Materiais educativos impressos (MEI) elaborados pelo Sistema Único de Saúde (SUS), como cartazes, folhetos e cartilhas, servem para veicular informações sobre a esquistossomose em ambientes de grade circulação de pessoas (unidades de saúde, escolas, estabelecimentos comerciais) ou para serem diretamente distribuídos à comunidade (MASSARA C. L.; MURTA, 2016; BRASIL, 2008).

## 2.5 Epidemiologia espacial

A epidemiologia espacial busca estimativas empíricas de relações entre indicadores epidemiológicos da ocorrência de doenças e variáveis ambientais e/ou socioeconômicas, que geralmente são modeladas como covariáveis. A adição das covariáveis tem o objetivo de promover o entendimento dos fatores que influenciam a distribuição espacial da doença, fornecer bases para explicações causais, acompanhar a disseminação de agravos à saúde, fazer previsões espaciais para intervenções geograficamente orientadas, e indicar os riscos aos quais uma população está exposta (HINO et al., 2006; HAMM et al., 2015).

A epidemiologia tem recorrido à aplicação de dados de observação da Terra com objetivos amplos, expandindo o uso dos SIGs e incorporando variáveis ambientais que influenciam a distribuição de doenças infecciosas, bem como caracterizam o habitat, que pode mudar de acordo com o clima e conforme o uso e cobertura da Terra se

alteram, de patógenos, vetores e hospedeiros (HAMM et al., 2015). Além disso, Bailey (2001) afirma que o aumento do interesse no estudo da distribuição geográfica da incidência de doenças e sua relação com potenciais fatores de risco promoveu o desenvolvimento de técnicas e métodos estatísticos apropriados para problemas em epidemiologia espacial, contendo algumas ferramentas poderosas, versáteis e relevantes.

## 2.6 GAM

O Modelo Aditivo Generalizado (GAM) é entendido como uma extensão do Modelo Linear Generalizado (GLM), de onde herda a interpretabilidade dos resultados. No entanto, é muito mais flexível por ser não paramétrico, ou seja, não possuir dados ou população com estrutura fixa, o que o aproxima das técnicas de machine learning. Funções de previsão mais flexíveis podem descobrir padrões ocultos nos dados. Quando esse tipo de flexibilidade é aplicada nos modelos lineares, corre-se o risco de causar overfitting, que ocorre nas situações em que o modelo se adaptou muito bem aos dados de treinamento, mas não generaliza bem para novos dados porque não aprendeu de fato o que diferencia aqueles dados (LARSEN, 2015).

O GAM é uma técnica de modelagem aditiva onde o impacto das variáveis preditivas é capturado por meio de funções suaves ( $s_n(x_n)$ ), equivalentes aos parâmetros  $\beta$  nos modelos lineares. Essas funções não precisam ser lineares, podem ser desconhecidas, e não paramétricas. Posteriormente adiciona-se os termos ao modelo para prever o valor da função de ligação  $g()$ , que equivale à esperança da variável resposta. O ajuste da suavidade no GAM é importante porque aproxima o resultado da ideia de que relacionamentos preditivos são inerentemente suaves na natureza, embora o conjunto de dados possa sugerir uma relação mais ruidosa. A suavização pode ser controlada com a penalidade aplicada ao modelo através do valor de  $\lambda$ , servindo para compensação de viés/variância. Esse valor pode ser definido pelo analista ou estimado a partir dos dados (LARSEN, 2015).

Os graus de liberdade efetivos no GAM, quando não há penalização, equivalem ao número de parâmetros a serem estimados. Quando a penalização é aplicada, os graus de liberdade efetivos reduzem conforme o  $\lambda$  aumenta. O número de funções internas é dado por  $K$ , que pode ser calculado a partir de nós, que por sua vez, são valores de  $x_n$  ordenados e utilizados como referência para a construção da equação e curva do modelo. A equação geral do GAM é apresentada em 2.1 (LARSEN, 2015).

$$g(E(Y)) = \alpha + s_1(x_1) + \dots + s_p(x_p) \quad (2.1)$$

## 2.7 Curva ROC

A curva ROC é uma forma gráfica de avaliação da qualidade de um modelo que é construída tendo no seu eixo horizontal os valores de especificidade (verdadeiro negativo), ou seja, a proporção de observações caracterizadas pela não ocorrência do evento que foram classificadas como não portadoras da característica, e no eixo vertical os valores de sensibilidade (verdadeiro positivo), isto é, a proporção de observações caracterizados pela ocorrência do evento que foram classificadas como portador da característica. O modelo é avaliado através da comparação dos resultados da sua classificação com uma classificação aleatória, que no gráfico equivale a uma reta de 45° partindo da origem até o canto superior direito. Assim, melhor é o desempenho do modelo quanto mais distante a curva estiver da diagonal principal do gráfico, ou seja, quanto maior for a área entre a curva ROC produzida e a diagonal principal, melhor é o desempenho global do modelo (GARCÍA J. P.; FERREIRA, 2021).

### 3 MATERIAIS E MÉTODOS

#### 3.1 Área de estudo

No estado de São Paulo, o coeficiente de incidência de casos novos de esquistossomose por 100 mil habitantes foi de 0,07 em 2020. Foram registradas 145 notificações da doença em 63 município de residência diferentes. Embora haja redução nos casos da infecção desde 2013, São Paulo foi o estado com o maior número de casos de esquistossomose em 2020 (SECRETARIA DE ESTADO DE SAÚDE DE SÃO PAULO, 2021). A área considerada nesse estudo foi a Região Geográfica Imediata (RGI) de Ourinhos, no estado de São Paulo.

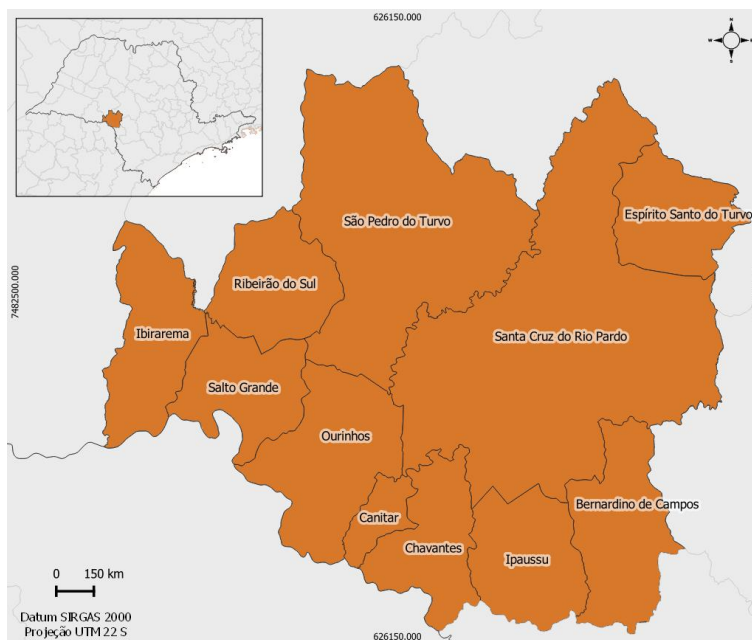
A divisão em RGIs realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE) em 2017 é vinculada aos processos sociais, políticos e econômicos brasileiros e atualiza a antiga divisão regional do Brasil em microrregiões realizada na década de 1990. As RGIs são embasadas na rede urbana do país e consistem em regiões estruturadas a partir de centros urbanos próximos para a satisfação das necessidades imediatas das populações, como compras de bens de consumo duráveis e não duráveis; busca de trabalho; procura por serviços de saúde e educação; e prestação de serviços públicos, entre outros (IBGE, s.d.).

A RGI de Ourinhos (Figura ??) está inserida na Bacia Hidrográfica do Médio Paranapanema e inclui 11 municípios (Ourinhos, Bernardino de Campos, Canitar, Chavantes, Espírito Santo do Turvo, Ibirarema, Ribeirão do Sul, São Pedro do Turvo, Santa Cruz do Rio Pardo, Salto Grande e Ipaussu). A região apresenta grande disponibilidade de água superficial e subterrânea, com presença de rios e lagos. A vegetação natural remanescente divide espaço com a atividade agrícola nas áreas rurais, onde há cultivo de cana-de-açúcar, uva, milho, soja, pêssego, e outros (SIGRH, s.d.). A Bacia do Médio Paranapanema é considerada uma importante área endêmica da esquistossomose e concentra aspectos favoráveis para a expansão da doença: urbanização precária em termos de saneamento básico (PALASIO et al., 2021b), uso da terra para atividade agrícola, inclusive com pivôs de irrigação central, presença dos três hospedeiros intermediários do *S. mansoni* nos corpos de água doce e municípios com histórico de casos de esquistossomose (PALASIO et al., 2021a).

#### 3.2 Metodologia

A metodologia da análise de potencial presença de caramujos *Biomphalaria* com o uso de GAM pode ser dividida em 3 etapas: (i) definição, criação e organização das

Figura 3.1 - Mapa de localização da Região Geográfica Imediata (RGI) de Ourinhos



Fonte: Produção da autora.

variáveis; (ii) ajuste do modelo aditivo generalizado; e (iii) avaliação e discussão dos resultados, que será apresentada no Capítulo 4.

### 3.3 Definição, criação e organização das variáveis

Considerando-se o ciclo de vida do caramujo e a literatura de malacologia que aponta a dependência de disponibilidade de água doce lântica para a fixação e sobrevivência dos moluscos *Biomphalaria* (BRASIL, 2008; MOURA; AL., 2005; CARVALHO O. S.; CALDEIRA, 2004), escolheu-se para compor o banco de dados variáveis que indicam a presença de água no ambiente, além da presença presumida do caramujo. A distribuição espacial das variáveis é mostrada na Figura 3.2. A unidade espacial da análise foram células quadradas regulares de 500m de lado, que cobrem toda a área de estudo somando 16.355 unidades. Deste total, 30% (4.907) foi reservada para a validação do modelo com a curva ROC e 11.448 dados compuseram o banco de dados da análise.

Os dados de presença presumida de caramujos foram gerados a partir de classificação anterior das células da grade regular com potencial ou não para a existência de habitats dos caramujos. Inicialmente, a partir de revisão da literatura, criou-se uma

tipologia associada à presença de caramujos do gênero *Biomphalaria* para basear a classificação da paisagem relacionada ao hospedeiro intermediário da esquistossomose identificando as características do ambiente que o tornariam propenso ou não para o estabelecimento de habitats. A classificação da paisagem foi feita com o toolbox GeoDMA 2.0.3 do software TerraView, utilizando o classificador C5.0 Decision Tree. O plugin extrai as propriedades do objeto e usa técnicas de mineração de dados para construir de forma automática árvores de decisão para classificação de objetos (KORTING, 2012). Inicialmente coletou-se manualmente amostras de treinamento do classificador e 200 árvores de decisão foram montadas com o auxílio da opção "boosting" da ferramenta. A classificação do GeoDMA designou para cada célula a classe mais frequentemente atribuída por todas as árvores de decisão. Amostras de teste foram usadas para o cálculo da acurácia e exatidão global, de 91,8%, da classificação. Ao final do processo, cada célula foi classificada com um dos oito tipos de paisagem descritos na região. Em seguida, os dados de classe foram transformados em dados binários considerando-se o potencial positivo ou negativo para o estabelecimento de habitats do caramujo. Os dados foram divididos em dois bancos de dados: um para a aplicação do modelo, com 70% dos dados, e outro para a validação com a curva ROC, com 30% dos registros. A presença presumida de caramujos no banco de dados principal totalizou 8.791 entradas que assumiram valor "1", e 2.657 dados foram expressos pelo valor zero, representando não-ocorrência presumida. Nos dados de validação, houve 3.802 registros de ocorrência e 1.105 de não-ocorrência.

Os dados de uso e cobertura da terra de 2020 em formato matricial com resolução de 30m X 30m tiveram como fonte a Coleção 5 do MapBiomas. Houve agregação das classes de "formação vegetal", "floresta plantada", "soja", "cana", "lavoura perene", "outras lavouras temporárias" e "mosaico de agricultura e pastagem". Calculou-se a área de cada célula ocupada pela cobertura e o resultado compôs a variável de porcentagem de área vegetada.

A drenagem da RGI de Ourinhos foi extraída a partir de dados de elevação da missão SRTM de 2000, que têm 30 metros de resolução e são disponibilizados em formato matricial pelo Instituto de Pesquisa Geológica dos Estados Unidos (USGS) em seu website. Foi necessário realizar a união dos três arquivos que compreendem a área de estudo (S23W050, S23W051 e S24W050) para, em seguida, submeter a processamento na plataforma TerraHidro 5.1.1. Após testes e eliminação de sumidouros, o limiar 900 foi escolhido para a extração da drenagem em formato vetorial e para o cálculo da distância vertical até a drenagem com o uso do algoritmo Height Above the Nearest Drainage (HAND). O resultado matricial do HAND foi fatiado em duas

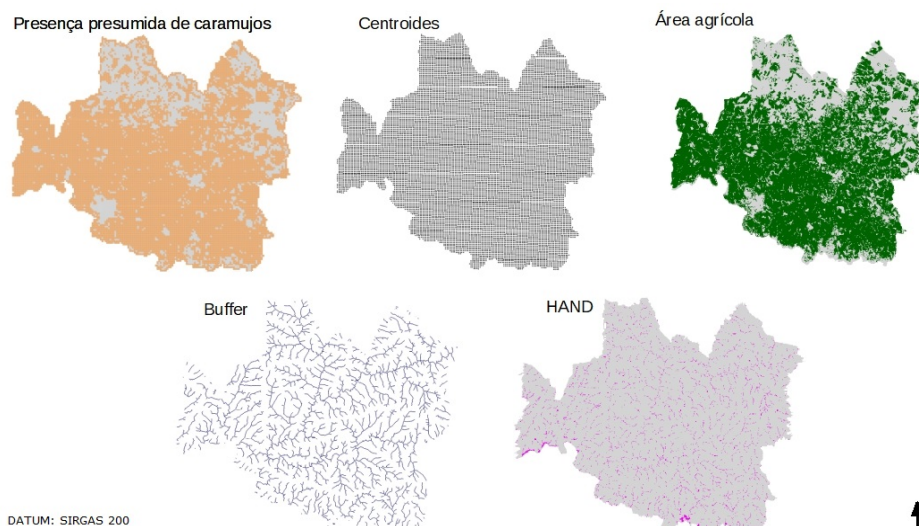


classes: até 1m e mais de 1m e foi incluída no banco de dados como uma das variáveis explicativas. A quarta variável explicativa foi a área de amortecimento (buffer) de 5m, produzido usando-se a drenagem extraída. As células que compreendiam o buffer receberam valor "1".

Tabela 3.1 - Variáveis incluídas no GAM.

<b>Dado</b>	<b>Fonte</b>	<b>Tipo</b>
Presença presumida de caramujos <i>Bi-omphalaria</i>	Classificação por árvore de decisão	Variável resposta quantitativa binária
Coordenadas	Centroides das células	Variável explicativa quantitativa contínua
Porcentagem de área agrícola	Coleção 5, Mapbiomas	Variável explicativa quantitativa contínua
HAND	Geração a partir da drenagem	Variável explicativa quantitativa binomial
Buffer da drenagem	Geração a partir da drenagem	Variável explicativa quantitativa binária

Figura 3.2 - Distribuição espacial dos dados considerados na análise.



Fonte: MAPBIOMAS, 2020. Produção da autora.

### 3.4 Ajuste do Modelo

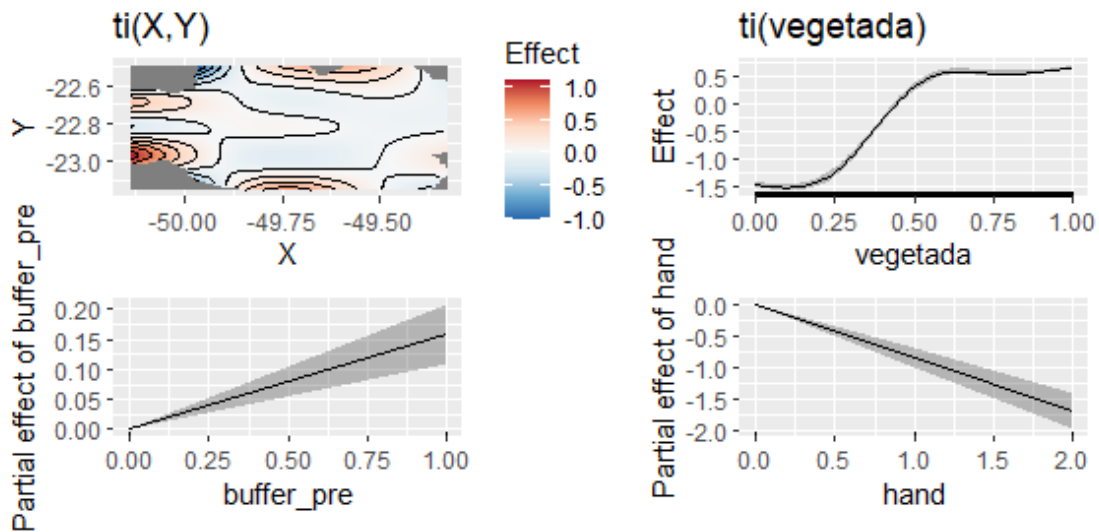
As análises foram realizadas no software estatístico R Studio versão 1.2.5033 e R versão 3.6, com os pacotes MGCV, para ajuste do modelo; Gratia e Ggplot2, para personalização da aparência dos gráficos e mapa; e ROCR para geração da curva ROC. Definiu-se a família "Betar" e função de ligação "Logit" por serem as adequadas para variável resposta binária. O método REML (Restricted Maximum Likelihood) foi utilizado para a estimação. Realizou-se testes com a adição, retirada e suavização das variáveis até obter-se o menor valor de AIC, maior  $R^2$  ajustado e p-valor abaixo de 0,05 para as variáveis e funções de suavização incluídas. A equação final do modelo, apresentada em 4.1, suavizou as variáveis X, Y e área vegetada com a função "ti".

## 4 ANÁLISE E RESULTADOS

A equação que descreve a presença potencial de caramujos *Biomphalaria* é especificada em 4.1. O  $R^2$  ajustado indica que 50,2% da presença potencial de caramujos pode ser explicada pelo modelo, que tem os resultados apresentados na Figura 4.1. Todas as variáveis e suavizações foram significativas considerando-se o intervalo de confiança de 95%.

$$vresp \sim ti(X, Y) + ti(vegetada) + buffer + hand, \quad \text{Sendo } ti \text{ função de suavização} \quad (4.1)$$

Figura 4.1 - Resultado do modelo.

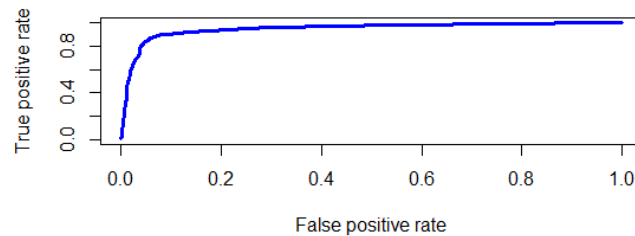


Fonte: Produção da autora.

O valor  $K$  foi definido pelo software e mostrou-se adequado para captar o comportamento dos dados, o que pode ser observado pela proximidade dos números de  $K'$  e os Graus de Liberdade Efetivos (EDF): 16 e 13,55, respectivamente, para a suavização de  $X$  e  $Y$ ; e 4 e 3,98 para a área vegetada. Parte da equação de penalização, o valor  $\lambda$  não variou no teste com métodos diferentes, mostrando boa estimação. O teste indicou que o modelo não apresenta concurvidade e a curva ROC (Figura 4.2), calculada com 30% (4.907 registros) do banco de dados original, mostrou que o modelo

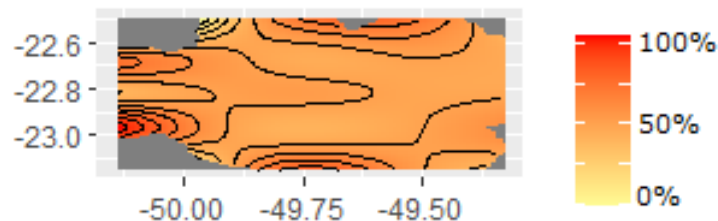
está acima do desempenho de uma classificação aleatória.

Figura 4.2 - Gráfico da curva ROC calculada para o modelo



Fonte: Produção da autora.

Figura 4.3 - Mapa de potencial presença de caramujos *Biomphalaria*



Fonte: Produção da autora.

O resultado do modelo apresenta suavização significativa da variável de área vegetada, ao mesmo tempo que mantém as relações lineares dos dados de Hand e Buffer, que não puderam ser suavizados por serem variáveis binomiais. O mapa da Figura 4.3 identifica a variação potencial, em porcentagem, de presença de caramujos na área de estudo. A suavização é perceptível quando comparada com os dados de entrada da variável resposta, que apresentava como característica oriunda da classificação por árvores de decisão em grade celular regular transições mais rígidas entre as áreas de presença e não-presença de caramujo. É possível perceber certa concordância nas partes com maior e menor valores, especialmente nas bordas. O GAM permite que a componente espacial do modelo seja suavizada com diferentes funções, inclusive de forma anisotrópica. O resultado final, no entanto, carrega as incertezas das va-

riáveis e está condicionado às limitações da variável resposta, que foi classificada utilizando-se algumas das mesmas variáveis usada no GAM.

## 5 CONCLUSÕES

Neste trabalho foi utilizado o Modelo Aditivo Generalizado para analisar a potencial presença de caramujos *Biomphalaria* na RGI de Ourinhos, São Paulo. O modelo apresenta resultados satisfatórios por ser flexível e considerar que relacionamentos preditivos são inerentemente suaves na natureza. O resultado também aponta a concordância de áreas de alto e baixo potencial de presença dos caramujos, especialmente nas bordas da área de estudo, com os dados originais da variável resposta. O GAM pode ser usado associando-se a componente espacial com outras variáveis que também ajudam a explicar a presença dos caramujos, como a disponibilidade de água no ambiente. A parte espacial pode ser suavizada com diferentes funções, sendo suportada inclusive forma anisotrópica.

Reitera-se que a identificação de áreas de potencial presença dos caramujos hospedeiros intermediários da esquistossomose é essencial para o planejamento de intervenções de controle efetivas. Ações como aplicação de moluscidas e outras intervenções físicas em focos e criadouros acontecem em áreas de extensão restritas. A correta identificação dessas áreas com o uso do GAM pode orientar a organização de trabalhos em campo, tornando as ações mais diretas e eficazes. Medidas de saneamento ambiental, que também fazem parte das ações de controle da doença, podem ter planejamentos beneficiados com a disponibilidade de dados de áreas de potencial presença do hospedeiro intermediário.

Por fim, como perspectiva futura para o trabalho é o uso de dados de criadouros dos caramujos *Biomphalaria*, obtidos em campo, para refinar a variável resposta do modelo.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ANARUMA, F.; SANTOS, F.; SANTANA, J. M. **Indicadores da relação entre estrutura da paisagem, degradação ambiental e esquistossomose mansoni**. Caxambu, 2007. 5
- ANDRADE, Z. A. Schistosomal hepatopathy. **Memórias do Instituto Oswaldo Cruz**, v. 99, n. 1, p. 5157, 2004. 4
- BAILEY, T. C. Spatial statistical methods in health. **Cad. Saúde Pública**, v. 17, n. 5, 2001. <https://doi.org/10.1590/S0102-311X2001000500011>. 8
- BRASIL, M. S. **Vigilância e controle de moluscos de importância epidemiológica: diretrizes: Programa de Vigilância e Controle da Esquistossomose**. Brasília: Editora do MS, 2008. 3, 4, 6, 7, 11
- CARVALHO O. S.; CALDEIRA, R. L. **Identificação morfológica de Biomphalaria glabrata, B. tenagophila e B. straminea hospedeiros intermediários do Schistosoma mansoni**. Belo Horizonte: Centro de Pesquisas René Rachou/FIOCRUZ, 2004. 5, 11
- CARVALHO, O. S. e. a. **Moluscos Brasileiros de Importância Médica**. Belo Horizonte: Fiocruz/Centro de Pesquisas René Rachou, 2014. 1, 4
- CASTAGNA, C. L. **Medida da vulnerabilidade ambiental para esquistossomose mansoni em função da estrutura da paisagem**. 99 p. Dissertação ao (Mestrado em Engenharia Civil) — Universidade Estadual de Campinas (Unicamp), Campinas, 2010. 1, 3, 4, 5, 6
- CHIEFFI, P. P.; WALDMAN, E. A. Aspectos particulares do comportamento epidemiológico da esquistossomose mansônica no estado de são paulo, brasil. **Cadernos de Saúde Pública**, v. 4, n. 4, p. 257–275, 1988. 3
- CONFALONIERI, U. E. C. Saúde na amazônia: um modelo conceitual para a análise de paisagens e doenças. **Estud av.**, v. 19, n. 53, p. 221–236, 2005. 1
- FIOCRUZ. **Esquistossomose**. Rio de Janeiro, 2014. Disponível em: <https://portal.fiocruz.br/taxonomia-geral-7-doencas-relacionadas/esquistossomose>. Acesso em: 21 mar 2021. 1
- GARCÍA J. P.; FERREIRA, J. C. P. C. M. Análise roc: uma aliada na pandemia. **Jornal Brasileiro de Pneumologia**, v. 47, n. 2, 2021. 9

HAMM, N. A. S.; MAGALHÃES, R. J. S.; CLEMENTS, A. C. A. Earth observation, spatial data quality, and neglected tropical diseases. **PLoS Negl Trop Dis**, v. 9, n. 12, 2015. (<https://doi.org/10.1371/journal.pntd.0004164>). 1, 7, 8

HINO, P.; VILLA, T. C. S.; SASSAKI, C. M.; NOGUEIRA, J. A.; SANTOS, C. B. Geoprocessamento aplicado à área da saúde. **Revista Latino-Americana de Enfermagem**, v. 14, n. 6, Nov-Dez 2006. 7

IBGE. **Divisão Regional do Brasil**. Rio de Janeiro, s.d. Disponível em: <<https://www.ibge.gov.br/geociencias/organizacao-do-territorio/divisao-regional/15778-divisoes-regionais-do-brasil.html>>. Acesso em: 01 set 2021. 10

KATZ, N. **Inquérito Nacional de Prevalência da Esquistossomose mansoni e Geo-helminthoses**. Belo Horizonte: Fiocruz/Centro de Pesquisas René Rachou, 2018. 1

KATZ, N.; PEIXOTO, S. V. Análise crítica da estimativa do número de portadores de esquistossomose no Brasil. **Rev. Soc. Bras. Med. Trop**, v. 31, n. 3, 2000. 3

KORTING, T. S. **GEODMA: A toolbox integrating data mining with object-based and multi-temporal analysis of satellite remotely sensed imagery**. 97 p. Tese (Tese de doutorado) — INPE, São José dos Campos, 2012. 12

LARSEN, K. Gam : The predictive modeling silver bullet. In: . [S.l.: s.n.], 2015. 8

MASSARA C. L.; MURTA, F. L. G. M. J. E. M. J. e. a. Caracterização de materiais educativos impressos sobre esquistossomose, utilizados para educação em saúde em áreas endêmicas no Brasil. **Epidemiol. Serv. Saúde**, v. 25, n. 3, 2016. 7

MOURA, A. C. M.; AL. et. **Atualização de mapa de drenagem como subsídio para montagem de SIG para a análise da distribuição da esquistossomose em Minas Gerais**. Goiânia, 2005. 6, 11

PALASIO, R. G. S.; AZEVEDO, T. S.; TUAN, R.; CHIARAVALLOTI-NETO, F. Modelling the present and future distribution of *Biomphalaria* species along the watershed of the middle paranapanema region, São Paulo, Brazil. **Acta Trop**, v. 214, 2021. 10

PALASIO, R. G. S.; BORTOLETO, A. N.; R., T.; F., C.-N. Geographic information system-based association between the sewage network, geographical



location of intermediate hosts, and autochthonous cases for the estimation of risk areas of schistosomiasis infection in ourinhos, são paulo, brazil. **Rev. Soc. Bras. Med. Trop.**, v. 54, 2021. 10

PESSOA, S. B.; MARTINS, A. V. **Parasitologia Médica**. Rio de Janeiro: Guanabara Koogan, 1988. 11<sup>a</sup> edição. 3

PIERI, O. S. Perspectivas no controle ambiental dos moluscos vetores da esquistossomose. In: BARBOSA, F. S. (Ed.). **Tópicos em Malacologia Médica**. Rio de Janeiro: Fiocruz, 1995. p. 239–252. 4

SAÚDE, B. M. da Saúde; Secretaria de Vigilância em. **Guia de vigilância epidemiológica**. Brasília: Ministério da Saúde, 2009. 1, 7

SECRETARIA DE ESTADO DE SAÚDE DE SÃO PAULO. **Casos Notificados por GVE, municípios de residência, infecção, ano**. São Paulo, 2021. Disponível em: <[https://www.saude.sp.gov.br/resources/cve-centro-de-vigilancia-epidemiologica/areas-de-vigilancia/doencas-de-transmissao-por-vetores-e-zoonoses/dados/esquisto/esquisto\\_dados.pdf](https://www.saude.sp.gov.br/resources/cve-centro-de-vigilancia-epidemiologica/areas-de-vigilancia/doencas-de-transmissao-por-vetores-e-zoonoses/dados/esquisto/esquisto_dados.pdf)>. Acesso em: 22 mar. 2021. 10

SIGRH. **Apresentação – Bacia do Médio Pranapanema**. Goiânia, s.d. Disponível em: <<http://www.sigrh.sp.gov.br/cbhmp/apresentacao>>. Acesso em: 04 abr 2021. 10

TELES, H. M. S. Distribuição geográfica das espécies dos caramujos transmissores de schistosoma mansoni no estado de são paulo. **Rev. Soc. Bras. Med. Trop.**, v. 38, n. 5, 2005. 5